

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:29:45 ; Search time 158 Seconds

(without alignments)
49.568 Million cell updates/sec

Title: US-09-674-794A-14

Perfect score: 112

Sequence: 1 GGGGGGGGGGGGGGGGGG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112	100.0	20	9	US-09-953-321-12
2	112	100.0	20	10	US-09-910-483-96
3	112	100.0	20	10	US-09-967-237-116
4	112	100.0	20	10	US-09-911-261A-21
5	112	100.0	20	10	US-09-969-748C-39
6	112	100.0	20	10	US-09-949-039-57
7	112	100.0	20	14	US-10-112-612-75
8	112	100.0	20	14	US-10-057-408-21
9	112	100.0	20	14	US-10-021-818-7
10	112	100.0	20	14	US-10-072-301-4
11	112	100.0	20	14	US-10-071-866-4
					Sequence 12, Appl
					Sequence 96, Appl
					Sequence 116, Appl
					Sequence 21, Appl
					Sequence 39, Appl
					Sequence 57, Appl
					Sequence 75, Appl
					Sequence 21, Appl
					Sequence 7, Appl
					Sequence 4, Appl

12	112	100.0	20	14	US-10-112-691-75	Sequence 75, Appl
13	112	100.0	20	14	US-10-287-941-1	Sequence 1, Appl
14	112	100.0	20	15	US-10-360-828-4	Sequence 4, Appl
15	112	100.0	20	15	US-10-032-0378-123	Sequence 123, App
16	112	100.0	20	15	US-10-029-988B-123	Sequence 123, App
17	112	100.0	20	15	US-10-032-423A-123	Sequence 123, App
18	112	100.0	20	15	US-10-622-108-30	Sequence 30, Appl
19	112	100.0	20	15	US-10-029-926B-123	Sequence 123, App
20	112	100.0	20	15	US-10-333-487-21	Sequence 21, Appl
21	112	100.0	20	16	US-10-475-540-9	Sequence 9, Appl
22	112	100.0	20	16	US-10-334-235-20	Sequence 20, Appl
23	112	100.0	20	16	US-10-746-149-29	Sequence 29, Appl
24	112	100.0	20	16	US-10-609-019-8	Sequence 8, Appl
25	112	100.0	20	16	US-10-470-987-20	Sequence 20, Appl
26	112	100.0	20	16	US-10-746-943-84	Sequence 84, Appl
27	112	100.0	20	16	US-10-492-729-16	Sequence 16, Appl
28	112	100.0	20	16	US-10-634-740-20	Sequence 20, Appl
29	112	100.0	20	17	US-10-885-223-149	Sequence 149, App
30	112	100.0	20	17	US-10-792-498-31	Sequence 31, Appl
31	112	100.0	20	17	US-10-770-304-5	Sequence 5, Appl
32	112	100.0	20	17	US-10-506-651-20	Sequence 20, Appl
33	112	100.0	20	17	US-10-954-094-149	Sequence 149, App
34	112	100.0	20	18	US-10-988-485-4	Sequence 4, Appl
35	112	100.0	20	18	US-10-792-682-75	Sequence 75, Appl
36	112	100.0	24	9	US-09-851-271A-12	Sequence 12, Appl
37	112	100.0	24	16	US-10-769-831-17	Sequence 17, Appl
38	112	100.0	24	17	US-10-770-140-17	Sequence 17, Appl
39	112	100.0	25	10	US-09-969-748C-40	Sequence 40, Appl
40	112	100.0	25	10	US-09-949-039-56	Sequence 56, Appl
41	112	100.0	25	13	US-10-081-281-32	Sequence 32, Appl
42	112	100.0	25	14	US-10-013-173-11	Sequence 11, Appl
43	112	100.0	25	14	US-10-150-762-11	Sequence 11, Appl
44	112	100.0	25	14	US-10-244-821-11	Sequence 11, Appl
45	112	100.0	25	14	US-10-261-798-91	Sequence 91, Appl
46	112	100.0	25	16	US-10-769-831-16	Sequence 16, Appl
47	112	100.0	25	16	US-10-634-740-21	Sequence 21, Appl
48	112	100.0	25	17	US-10-770-140-1	Sequence 1, Appl
49	112	100.0	25	17	US-10-770-140-16	Sequence 16, Appl
50	112	100.0	25	17	US-10-770-304-4	Sequence 4, Appl
51	112	100.0	25	17	US-10-506-651-14	Sequence 14, Appl
52	112	100.0	26	16	US-10-646-308-36	Sequence 36, Appl
53	112	100.0	30	10	US-09-833-203-19	Sequence 19, Appl
54	112	100.0	30	16	US-10-492-729-17	Sequence 17, Appl
55	112	100.0	30	16	US-10-634-740-22	Sequence 22, Appl
56	112	100.0	30	17	US-10-792-498-32	Sequence 32, Appl
57	112	100.0	35	14	US-10-013-173-30	Sequence 30, Appl
58	112	100.0	35	14	US-10-150-762-30	Sequence 30, Appl
59	112	100.0	35	14	US-10-244-821-30	Sequence 30, Appl
60	112	100.0	40	13	US-10-081-400-1	Sequence 1, Appl
61	112	100.0	40	14	US-10-005-438-9	Sequence 9, Appl
62	112	100.0	40	15	US-10-257-864A-137	Sequence 137, App
63	112	100.0	40	15	US-10-257-864A-138	Sequence 138, App
64	112	100.0	40	16	US-10-608-710-6	Sequence 6, Appl
65	112	100.0	40	16	US-10-768-873-1	Sequence 1, Appl
66	112	100.0	40	16	US-10-399-585-176	Sequence 176, App
67	112	100.0	40	16	US-10-399-585-177	Sequence 177, App
68	112	100.0	40	16	US-10-645-085A-137	Sequence 137, App
69	112	100.0	40	16	US-10-645-085A-138	Sequence 138, App
70	112	100.0	40	16	US-10-492-729-18	Sequence 18, Appl
71	112	100.0	40	17	US-10-917-899-9	Sequence 9, Appl
72	112	100.0	40	17	US-10-792-498-33	Sequence 33, Appl
73	112	100.0	40	20	US-11-016-189-1	Sequence 1, Appl
74	112	100.0	50	10	US-09-949-039-32	Sequence 32, Appl
75	112	100.0	50	17	US-10-842-054-15	Sequence 15, Appl
76	112	100.0	50	17	US-10-792-498-34	Sequence 34, Appl
77	112	100.0	50	18	US-10-841-819B-31	Sequence 31, Appl
78	112	100.0	60	9	US-09-832-297A-12	Sequence 12, Appl
79	112	100.0	60	10	US-09-833-203-18	Sequence 18, Appl
80	112	100.0	60	17	US-10-506-651-17	Sequence 17, Appl
81	112	100.0	80	17	US-10-842-054-10	Sequence 10, Appl
82	112	100.0	209	15	US-10-449-831A-206	Sequence 206, App
83	112	100.0	211	16	US-10-437-963-126553	Sequence 126553, App
84	112	100.0	233	9	US-09-815-837-95	Sequence 95, Appl

85	112	100.0	234	13	US-10-081-281-119	Sequence 119, App
86	112	100.0	239	15	US-10-449-831A-194	Sequence 194, App
87	112	100.0	242	15	US-10-449-831A-190	Sequence 190, App
88	112	100.0	247	15	US-10-620-278-21	Sequence 21, Appl
89	112	100.0	247	17	US-10-620-049-21	Sequence 21, Appl
90	112	100.0	250	14	US-10-072-301-21	Sequence 21, Appl
91	112	100.0	250	14	US-10-072-301-29	Sequence 29, Appl
92	112	100.0	250	14	US-10-071-866-21	Sequence 21, Appl
93	112	100.0	250	14	US-10-071-866-29	Sequence 29, Appl
94	112	100.0	250	15	US-10-360-828-21	Sequence 21, Appl
95	112	100.0	250	15	US-10-360-828-29	Sequence 29, Appl
96	112	100.0	250	18	US-10-988-485-21	Sequence 21, Appl
97	112	100.0	250	18	US-10-988-485-29	Sequence 29, Appl
98	112	100.0	252	9	US-09-815-837-91	Sequence 91, Appl
99	112	100.0	253	14	US-10-072-301-17	Sequence 17, Appl
100	112	100.0	253	14	US-10-072-301-23	Sequence 23, Appl

ALIGNMENTS

```

RESULT 1
US-09-953-321-12
; Sequence 12, Application US/09953321
; Patent No. US20020115083A1
; GENERAL INFORMATION:
; APPLICANT: PLUCKTHUN, ANDREAS
; APPLICANT: HANES, JOZEF
; APPLICANT: JERMUTUS, LUTZ
; TITLE OF INVENTION: NOVEL METHODS FOR IDENTIFYING NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: (POLY) PEPTIDES THAT INTERACT WITH TARGET MOLECULES
; FILE REFERENCE: PLUCK/1 CON2
; CURRENT APPLICATION NUMBER: US/09/953,321
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 09/425,585
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide Linker
US-09-953-321-12

```

```
Query Match      100.0%; Score 112; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy

1	GGGSGGGSGGGSGGGGS	20

Dβ

1	GGGSGGGSGGGSGGGGS	20
---	-------------------	----

```

RESULT 2
US-09-1910-483-96
; Sequence 96, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; APPLICANT: REMO, JOHN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

;
;
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Linker
; OTHER INFORMATION: peptide
US-09-910-483-96

```

Query Match	100.0%;	Score 112;	DB 10;	Length 20;
Best Local Similarity	100.0%;	Pred. No. 0.0018;		
Matches 20: Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy

1	G G G S G G G S G G G S G G G S	20

pB

1	G G G S G G G S G G G S G G G S	20

```

RESULT 3
US-09-967-237-116
; Sequence 116, Application US/09967237
; Publication No. US2003004982A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.58-2
; CURRENT APPLICATION NUMBER: US/09/967,237
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/178,115
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 116
; LENGTH: 20
; TYPE: PRT
; ORGANISM: HUMAN
US-09-967-237-116

```

```
Query Match      100.0%; Score 112; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy

1	G G G G C C G G S G C C S G G G S	20

Dd

1	G G G G C C G G S G C C S G G G S	20

```

RESULT 4
US-09-911-261A-21
; Sequence 21, Application US/09911261A
; Publication No. US20030134350A1
; GENERAL INFORMATION:
; APPLICANT: Sera, Takashi
; TITLE OF INVENTION: Zinc Finger Domain Recognition Code and Uses Thereof
; FILE REFERENCE: 109845.135
; CURRENT APPLICATION NUMBER: US/09/911,261A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 60/220,060
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flexible linker
US-09-911-261A-21

```

Query Match	100.0%;	Score 112;	DB 10;	Length 20;
Best Local Similarity	100.0%;	Pred. No. 0.0018;		
Matches 20: Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20

Db	1	GGGGGGGGGGGGGGGGG	20		Db	1	GGGGGGGGGGGGGGGGG	20	
RESULT 5									
US-09-969-748C-39									
; Sequence 39, Application US/09969748C									
; Publication No. US20030161809A1									
; GENERAL INFORMATION:									
; APPLICANT: ARIZEKE PHARMACEUTICALS, INC.									
; APPLICANT: HOUSTON, Lou, L.									
; APPLICANT: SHERIDAN, Philip, J.									
; APPLICANT: HAWLEY, Stephen									
; APPLICANT: GLYNN, Jacqueline, M.									
; APPLICANT: CHAPIN, Steven									
; APPLICANT: BASU, Amresh									
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE									
; TITLE OF INVENTION: AGENTS ACROSS CELLULAR BARRIERS									
; FILE REFERENCE: 057220-0303									
; CURRENT APPLICATION NUMBER: US/09/969,748C									
; CURRENT FILING DATE: 2002-12-10									
; PRIOR APPLICATION NUMBER: US 60/267,601									
; PRIOR FILING DATE: 2001-02-09									
; PRIOR APPLICATION NUMBER: US 60/248,819									
; PRIOR FILING DATE: 2000-11-14									
; PRIOR APPLICATION NUMBER: US 60/248,478									
; PRIOR FILING DATE: 2000-11-13									
; PRIOR APPLICATION NUMBER: US 60/237,929									
; PRIOR FILING DATE: 2000-10-02									
; NUMBER OF SEQ ID NOS: 115									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 39									
; LENGTH: 20									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: linker sequence									
US-09-969-748C-39									
Query Match 100.0%; Score 112; DB 10; Length 20;									
Best Local Similarity 100.0%; Pred. No. 0.0018;									
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GGGGGGGGGGGGGGGGG	20		Qy	1	GGGGGGGGGGGGGGGGG	20	
Db	1	GGGGGGGGGGGGGGGGG	20		Db	1	GGGGGGGGGGGGGGGGG	20	
RESULT 6									
US-09-949-039-57									
; Sequence 57, Application US/09949039									
; Publication No. US20030166160A1									
; GENERAL INFORMATION:									
; APPLICANT: HAWLEY, STEPHEN B.									
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE									
; TITLE OF INVENTION: BINDING REGIONS DIRECTED TO TRANSCYTOTIC LIGANDS									
; FILE REFERENCE: 057220/1301									
; CURRENT APPLICATION NUMBER: US/09/949,039									
; CURRENT FILING DATE: 2001-09-06									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 57									
; LENGTH: 20									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Gly-Ser linker									
US-09-949-039-57									
Query Match 100.0%; Score 112; DB 10; Length 20;									
Best Local Similarity 100.0%; Pred. No. 0.0018;									
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GGGGGGGGGGGGGGGGG	20		Qy	1	GGGGGGGGGGGGGGGGG	20	
Db	1	GGGGGGGGGGGGGGGGG	20		Db	1	GGGGGGGGGGGGGGGGG	20	
RESULT 7									
US-10-112-612-75									
; Sequence 75, Application US/10112612									
; Publication No. US20030027213A1									
; GENERAL INFORMATION:									
; APPLICANT: Zhu, Li									
; APPLICANT: Hua, Shaobing									
; TITLE OF INVENTION: HIGH THROUGHPUT GENERATION AND SCREENING OF FULLY HUMAN ANTIBODY									
; TITLE OF INVENTION: REPERTOIRE IN YEAST									
; FILE REFERENCE: 25636-719									
; CURRENT APPLICATION NUMBER: US/10/112,612									
; CURRENT FILING DATE: 2002-03-27									
; PRIOR APPLICATION NUMBER: US 09/603,663									
; PRIOR FILING DATE: 2000-06-23									
; NUMBER OF SEQ ID NOS: 75									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 75									
; LENGTH: 20									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Linker peptide									
US-10-112-612-75									
Query Match 100.0%; Score 112; DB 14; Length 20;									
Best Local Similarity 100.0%; Pred. No. 0.0018;									
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GGGGGGGGGGGGGGGGG	20		Qy	1	GGGGGGGGGGGGGGGGG	20	
Db	1	GGGGGGGGGGGGGGGGG	20		Db	1	GGGGGGGGGGGGGGGGG	20	
RESULT 8									
US-10-057-408-21									
; Sequence 21, Application US/10057408									
; Publication No. US20030082561A1									
; GENERAL INFORMATION:									
; APPLICANT: Sera, Takashi									
; TITLE OF INVENTION: Zinc Finger Domain Recognition Code and Uses Thereof									
; FILE REFERENCE: 109845.135									
; CURRENT APPLICATION NUMBER: US/10/057,408									
; CURRENT FILING DATE: 2002-01-23									
; PRIOR APPLICATION NUMBER: US 60/220,060									
; PRIOR FILING DATE: 2000-07-21									
; NUMBER OF SEQ ID NOS: 69									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 21									
; LENGTH: 20									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Flexible linker									
US-10-057-408-21									
Query Match 100.0%; Score 112; DB 14; Length 20;									
Best Local Similarity 100.0%; Pred. No. 0.0018;									
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GGGGGGGGGGGGGGGGG	20		Qy	1	GGGGGGGGGGGGGGGGG	20	
Db	1	GGGGGGGGGGGGGGGGG	20		Db	1	GGGGGGGGGGGGGGGGG	20	
RESULT 9									
US-10-021-818-7									
; Sequence 7, Application US/10021818									
; Publication No. US20030108566A1									
; GENERAL INFORMATION:									

```
; APPLICANT: Davis, Ronald W.
; APPLICANT: Vaillancourt, Peter
; TITLE OF INVENTION: Dimeric Fluorescent Polypeptides
; FILE REFERENCE: 25436/1652
; CURRENT APPLICATION NUMBER: US/10/021.818
; CURRENT FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: US 60/256,121
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic linker peptide
US-10-021-818-7

Query Match          100.0%; Score 112; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
   |||||
Db 1 GGGSGGGSGGGSGGGGS 20
   |||||

RESULT 10
US-10-072-301-4
; Sequence 4, Application US/10072301
; Publication No. US20030152913A1
; GENERAL INFORMATION:
; APPLICANT: Hua, Shao-bing
; APPLICANT: Pauling, Michelle H.
; APPLICANT: Zhu, Li
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 25636-718
; CURRENT APPLICATION NUMBER: US/10/072.301
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: G4S Linker
US-10-072-301-4

Query Match          100.0%; Score 112; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
   |||||
Db 1 GGGSGGGSGGGSGGGGS 20
   |||||

RESULT 11
US-10-071-866-4
; Sequence 4, Application US/10071866
; Publication No. US20030165988A1
; GENERAL INFORMATION:
; APPLICANT: Hua, Shao-bing
; APPLICANT: Pauling, Michelle H.
; APPLICANT: Zhu, Li
; TITLE OF INVENTION: HIGH THROUGHPUT GENERATION OF HUMAN MONOCLONAL ANTIBODY AGAINST P
; FILE REFERENCE: 25636-717
; CURRENT APPLICATION NUMBER: US/10/071.866
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 54
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: G4S Linker
US-10-071-866-4

Query Match          100.0%; Score 112; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
   |||||
Db 1 GGGSGGGSGGGSGGGGS 20
   |||||

RESULT 12
US-10-112-691-75
; Sequence 75, Application US/10112691
; Publication No. US20030165990A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing
; TITLE OF INVENTION: GENERATION OF HIGHLY DIVERSE LIBRARY OF EXPRESSION VECTORS VIA HO
; FILE REFERENCE: 25636-720
; CURRENT APPLICATION NUMBER: US/10/112,691
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/602,373
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-10-112-691-75

Query Match          100.0%; Score 112; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
   |||||
Db 1 GGGSGGGSGGGSGGGGS 20
   |||||

RESULT 13
US-10-287-941-1
; Sequence 1, Application US/10287941
; Publication No. US20030171552A1
; GENERAL INFORMATION:
; APPLICANT: Weidanz, Jon A.
; APPLICANT: Card, Kimberlyn F.
; APPLICANT: Sherman, Linda A.
; APPLICANT: Klinman, No. US20030171552Alman
; APPLICANT: Wong, Hing C.
; TITLE OF INVENTION: POLYSPECIFIC BINDING MOLECULES AND USES THEREOF
; FILE REFERENCE: 48531
; CURRENT APPLICATION NUMBER: US/10/287,941
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US/09/422,375
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
```



```
US-10-287-941-1
Query Match      100.0%; Score 112; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 14
US-10-360-828-4
; Sequence 4, Application US/10360828
; Publication No. US20030206909A1
; GENERAL INFORMATION:
; APPLICANT: Hua, Shaobing
; APPLICANT: Pauling, Michelle H.
; APPLICANT: Zhu, Li
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES AGAINST MEMBRANE PROTEINS
; FILE REFERENCE: 25636-727
; CURRENT APPLICATION NUMBER: US/10/360,828
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 10/071,866
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 10/072,301
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 10/133,978
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GAS Linker
US-10-360-828-4

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 15
US-10-032-037B-123
; Sequence 123, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-123

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 16
US-10-029-988B-123
; Sequence 123, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-123

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 17
US-10-032-423A-123
; Sequence 123, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-123

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 18
US-10-622-108-30
; Sequence 30, Application US/10622108
; Publication No. US20040063912A1
; GENERAL INFORMATION:
; APPLICANT: Blumberg, Richard S.
; APPLICANT: Lencer, Wayne I.
; APPLICANT: Simister, Neil E.
```

```
; APPLICANT: Bitonti, Alan J.
; TITLE OF INVENTION: CENTRAL AIRWAY ADMINISTRATION FOR SYSTEMIC DELIVERY OF THERAPEUTIC
; FILE REFERENCE: S01383.70011.US
; CURRENT APPLICATION NUMBER: US/10/622,108
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: US 10/435,608
; PRIOR FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: PCT/US02/21355
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/364,482
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
; US-10-622-108-30

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 19
US-10-029-926B-123
; Sequence 123, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-029-926B-123

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 20
US-10-333-487-21
; Sequence 21, Application US/10333487
; Publication No. US20040091878A1
; GENERAL INFORMATION:
; APPLICANT: Sera, Takahashi
; TITLE OF INVENTION: Zinc Finger Domain Recognition Code and Uses Thereof
; FILE REFERENCE: 109845-165
; CURRENT APPLICATION NUMBER: US/10/333,487
; CURRENT FILING DATE: 2003-01-17
; PRIOR APPLICATION NUMBER: PCT/EP01/08367
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/220,060
; PRIOR FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flexible linker
; US-10-333-487-21

Query Match      100.0%; Score 112; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 21
US-10-475-540-9
; Sequence 9, Application US/10475540
; Publication No. US20040110245A1
; GENERAL INFORMATION:
; APPLICANT: NAGAMUNE, Teruyuki
; APPLICANT: EIKEN KAGAKU KABUSHIKI KAISHA
; TITLE OF INVENTION: Method for an analysis using an interaction between
; FILE REFERENCE: PCT348EK
; CURRENT APPLICATION NUMBER: US/10/475,540
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: JP 2001-123866
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Flexible Linker
; US-10-475-540-9

Query Match      100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 22
US-10-334-235-20
; Sequence 20, Application US/10334235
; Publication No. US20040131591A1
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Ltd.
; APPLICANT: Kingsman, Alan
; APPLICANT: Bebbington, Christopher
; APPLICANT: Carroll, Miles
; APPLICANT: Ellard, Fiona
; APPLICANT: Kingsman, Susan
; APPLICANT: Myers, Kevin
; APPLICANT: Lamikandra, Abigail
; TITLE OF INVENTION: VECTOR SYSTEM
; FILE REFERENCE: 532682000920
; CURRENT APPLICATION NUMBER: US/10/334,235
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 10/060,585
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/GB00/04317
; PRIOR FILING DATE: 2000-11-13
```

; PRIOR APPLICATION NUMBER: US 09/445,375
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker
US-10-334-235-20

Query Match 100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGS 20
Db 1 GGGGGGGGGGGGGGGGGGS 20

RESULT 23
US-10-746-149-29
; Sequence 29, Application US/10746149
; Publication No. US20040172667A1
; GENERAL INFORMATION:
; APPLICANT: Cooper, Richard K.
; APPLICANT: Cadd, Gary G.
; TITLE OF INVENTION: Administration of Transposon-Based Vectors to Reproductive Organs
; FILE REFERENCE: 51687-0280 (51687-294923)
; CURRENT APPLICATION NUMBER: US/10/746,149
; PRIOR FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: US 60/392,415
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/441,381
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,447
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,405
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,502
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,392
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,377
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 10/609,019
; PRIOR FILING DATE: 2003-06-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-746-149-29

Query Match 100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGS 20
Db 1 GGGGGGGGGGGGGGGGGGS 20

RESULT 24
US-10-609-019-8
; Sequence 8, Application US/10609019
; Publication No. US20040197910A1
; GENERAL INFORMATION:

; APPLICANT: Cooper, Richard K.
; APPLICANT: Cadd, Gary G.
; APPLICANT: Fioretti, William C.
; APPLICANT: DeBoer, Kenneth F.
; TITLE OF INVENTION: Gene Regulation in Transgenic Animals Using a Transposon-Based
; FILE REFERENCE: 51687-0101 (51687-287015)
; CURRENT APPLICATION NUMBER: US/10/609,019
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/392,415
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/441,392
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,377
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,502
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,405
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,447
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,381
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-609-019-8

Query Match 100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGS 20
Db 1 GGGGGGGGGGGGGGGGGGS 20

RESULT 25
US-10-470-987-20
; Sequence 20, Application US/10470987
; Publication No. US20040219542A1
; GENERAL INFORMATION:
; APPLICANT: HOUSTON, LOU L.
; APPLICANT: SHERIDAN, PHILIP L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING, CHARACTERIZING,
; TITLE OF INVENTION: OPTIMIZING AND USING LIGANDS TO TRANSCYTOTIC MOLECULES
; FILE REFERENCE: 057220/0703
; CURRENT APPLICATION NUMBER: US/10/470,987
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: PCT/US02/03059
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/266,182
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 20
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Linker peptide
; FEATURE:
; NAME/KEY: REPEAT
; LOCATION: (1)..(5)
; OTHER INFORMATION: this region may be repeated 1-4 times
US-10-470-987-20

Query Match 100.0%; Score 112; DB 16; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.0018; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 26

US-10-746-943-84
; Sequence 84, Application US/10746943
; Publication No. US20040235011A1
; GENERAL INFORMATION:
; APPLICANT: Cooper, Richard K.
; APPLICANT: Fioretto, William C.
; APPLICANT: Cadd, Gary G.
; TITLE OF INVENTION: Production of Multimeric Proteins
; FILE REFERENCE: 51687-0230 (51687-294924)
; CURRENT APPLICATION NUMBER: US/10/746,943
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: US 60/392,415
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/441,381
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,447
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,405
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,502
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,377
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,392
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 10/609,019
; PRIOR FILING DATE: 2003-06-26
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 20
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-746-943-84

Query Match 100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 27

US-10-492-729-16
; Sequence 16, Application US/10492729
; Publication No. US20040259075A1
; GENERAL INFORMATION:
; APPLICANT: Dimitrov, Dimitar S
; APPLICANT: Moulard, Maxime
; APPLICANT: Xiao, Xiadong
; APPLICANT: Shu, Yuesi
; APPLICANT: Phogat, Sanjay K
; APPLICANT: Zhang, Mei-Yun
; APPLICANT: Burton, Dennis
; TITLE OF INVENTION: BROADLY CROSS-REACTIVE NEUTRALIZING ANTIBODIES AGAINST HUMAN
; FILE REFERENCE: 227062
; CURRENT APPLICATION NUMBER: US/10/492,729
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: PCT/US02/33165

; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: 60/329,709
; PRIOR FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-492-729-16

Query Match 100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 28

US-10-634-740-20
; Sequence 20, Application US/10634740
; Publication No. US20040265906A1
; GENERAL INFORMATION:
; APPLICANT: TING, ALICE
; TITLE OF INVENTION: GENETICALLY ENCODED FLUORESCENT REPORTERS OF KINASE,
; FILE REFERENCE: M0656.70097.US
; CURRENT APPLICATION NUMBER: US/10/634,740
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: US 60/425,578
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-10-634-740-20

Query Match 100.0%; Score 112; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 29

US-10-885-225-149
; Sequence 149, Application US/10885225
; Publication No. US20050009750A1
; GENERAL INFORMATION:
; APPLICANT: Sleeman, Matthew
; APPLICANT: Murison, J. Greg
; APPLICANT: Cao, Helen
; TITLE OF INVENTION: Fibroblast Growth Factor Receptors and Methods for Their Use
; FILE REFERENCE: 11000.1103
; CURRENT APPLICATION NUMBER: US/10/885,225
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: U.S. 60/484,877
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: U.S. 60/513,171
; PRIOR FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: U.S. 60/562,155
; PRIOR FILING DATE: 2004-04-04
; PRIOR APPLICATION NUMBER: U.S. 60/570,355

```

; PRIOR FILING DATE: 2004-05-12
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-10-885-225-149

Query Match      100.0%; Score 112; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGGGG 20

RESULT 30
US-10-792-498-31
; Sequence 31, Application US/10792498
; Publication No. US20050074865A1
; GENERAL INFORMATION:
; APPLICANT: Afeyan, Noubar B.
; APPLICANT: Lee, Frank D.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Das Gupta, Ruchira
; APPLICANT: Baynes, Brian
; TITLE OF INVENTION: ADZYMES AND USES THEREOF
; FILE REFERENCE: COTH-P03-001
; CURRENT APPLICATION NUMBER: US/10/792,498
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US 10/650,592
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US 60/406,517
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 60/423,754
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/430,001
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker
US-10-792-498-31

Query Match      100.0%; Score 112; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGGGG 20

RESULT 31
US-10-770-304-5
; Sequence 5, Application US/10770304
; Publication No. US20050096459A1
; GENERAL INFORMATION:
; APPLICANT: Schwabe, Nikolai
; TITLE OF INVENTION: OLIGOMERIC RECEPTOR LIGAND PAIR MEMBER COMPLEXES
; FILE REFERENCE: S-846
; CURRENT APPLICATION NUMBER: US/10/770,304
; CURRENT FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
US-10-770-304-5

; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-770-304-5

Query Match      100.0%; Score 112; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGGGG 20

RESULT 32
US-10-506-651-20
; Sequence 20, Application US/10506651
; Publication No. US20050106160A1
; GENERAL INFORMATION:
; APPLICANT: Dimitrov, Dimitar Stanchev
; APPLICANT: Chow, Yen-Hung
; APPLICANT: Phogat, Sanjay Kumar
; APPLICANT: Broder, Christopher Charles
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: HIV-1 Envelope Glycoproteins Stabilized by Flexible
; FILE REFERENCE: 015280-458000US
; CURRENT APPLICATION NUMBER: US/10/506,651
; CURRENT FILING DATE: 2004-09-02
; PRIOR APPLICATION NUMBER: WO PCT/US02/07144
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:exemplary
; OTHER INFORMATION: linker
US-10-506-651-20

Query Match      100.0%; Score 112; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGGGG 20

RESULT 33
US-10-954-094-149
; Sequence 149, Application US/10954094
; Publication No. US20050112642A1
; GENERAL INFORMATION:
; APPLICANT: Sleeman, Matthew
; APPLICANT: Murison, J. Greg
; APPLICANT: Cao, Zhihui
; TITLE OF INVENTION: Fibroblast Growth Factor Receptors and Methods for Their Use
; FILE REFERENCE: 11000.1037c8
; CURRENT APPLICATION NUMBER: US/10/954,094
; CURRENT FILING DATE: 2004-09-29
; PRIOR APPLICATION NUMBER: U.S. 10/157,444
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: U.S. 10/613,413
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: U.S. 60/484,877
US-10-954-094-149
```

; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: U.S. 60/513,171
; PRIOR FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: U.S. 60/562,155
; PRIOR FILING DATE: 2004-04-04
; PRIOR APPLICATION NUMBER: U.S. 60/570,355
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: U.S. 10/885,225
; PRIOR FILING DATE: 2004-07-04
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-10-954-094-149

Query Match 100.0%; Score 112; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 34
US-10-988-485-4
; Sequence 4, Application US/10988485
; Publication No. US20050123973A1
; GENERAL INFORMATION:
; APPLICANT: Hua, Shaobing
; APPLICANT: Pauling, Michelle H.
; APPLICANT: Zhu, Li
; TITLE OF INVENTION: METHODS FOR GENERATING MONOCLONAL ANTIBODY AGAINST FUSION PROTEIN
; FILE REFERENCE: 25636-722.301
; CURRENT APPLICATION NUMBER: US/10/988,485
; CURRENT FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US 10/071,866
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 10/072,301
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 10/133,978
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: G4S Linker
US-10-988-485-4

Query Match 100.0%; Score 112; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 35
US-10-792-682-75
; Sequence 75, Application US/10792682
; Publication No. US20050142562A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing

; TITLE OF INVENTION: HIGH THROUGHPUT GENERATION AND SCREENING OF FULLY HUMAN ANTIBODY
; TITLE OF INVENTION: REPERTOIRE IN YEAST
; FILE REFERENCE: 25636-719.301
; CURRENT APPLICATION NUMBER: US/10/792,682
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US 09/603,663
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 10/112,612
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker peptide
US-10-792-682-75

Query Match 100.0%; Score 112; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 36
US-09-851-271A-12
; Sequence 12, Application US/09851271A
; Patent No. US20020064824A1
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851,271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(24)
; OTHER INFORMATION: linker sequence followed by the stalling polypeptide sequence
US-09-851-271A-12

Query Match 100.0%; Score 112; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.002; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 37
US-10-769-831-17
; Sequence 17, Application US/10769831
; Publication No. US20040209295A1
; GENERAL INFORMATION:
; APPLICANT: Schwabe, Nikolai F
; APPLICANT: Tan, Linda C
; APPLICANT: Catherine, Napper E
; APPLICANT: Fry, Jeremy W
; APPLICANT: Pang, Susan

Tue Aug 23 10:23:59 2005

; TITLE OF INVENTION: CHIMERIC MHC PROTEIN AND OLIGOMER THEREOF
; FILE REFERENCE: S-844-US
; CURRENT APPLICATION NUMBER: US/10/769,831
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: PCT/EP03/09056
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-769-831-17

Query Match 100.0%; Score 112; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | |
Db 2 GGGSGGGSGGGSGGGGS 21
| | | | | | | | | | | | | | | | | | | | |

RESULT 38
US-10-770-140-17
; Sequence 17, Application US/10770140
; Publication No. US20050074848A1
; GENERAL INFORMATION:
; APPLICANT: Schwabe, Nikolai
; TITLE OF INVENTION: Chimeric Protein and Oligomer Thereof for Specific Targeting
; FILE REFERENCE: S-845
; CURRENT APPLICATION NUMBER: US/10/770,140
; CURRENT FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-770-140-17

Query Match 100.0%; Score 112; DB 17; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | |
Db 2 GGGSGGGSGGGSGGGGS 21
| | | | | | | | | | | | | | | | | | | | |

RESULT 39
US-09-969-748C-40
; Sequence 40, Application US/09969748C
; Publication No. US20030161809A1
; GENERAL INFORMATION:
; APPLICANT: ARIZEKE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, Lou, L.
; APPLICANT: SHERIDAN, Philip, J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GLYNN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven
; APPLICANT: BASU, Amaresh
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE
; FILE REFERENCE: 057220-0303
; CURRENT APPLICATION NUMBER: US/09/969,748C
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/267,601
; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: US 60/248,819
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/248,478
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/237,929
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: linker sequence
US-09-969-748C-40

Query Match 100.0%; Score 112; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | |

RESULT 40
US-09-949-039-56
; Sequence 56, Application US/09949039
; Publication No. US20030166160A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY, STEPHEN B.
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE
; BINDING REGIONS DIRECTED TO TRANSCYTOTIC LIGANDS
; FILE REFERENCE: 057220/1301
; CURRENT APPLICATION NUMBER: US/09/949,039
; CURRENT FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gly-Ser linker
US-09-949-039-56

Query Match 100.0%; Score 112; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | |

Search completed: August 19, 2005, 16:39:43
Job time : 159 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:26:09 ; Search time 39 Seconds
(without alignments)
49.342 Million cell updates/sec

Title: US-09-674-794A-14

Perfect score: 112

Sequence: 1 GGGGGGGGGGGGGGGGGG 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR.79.*

2: Pirl.*

3: Pirl2.*

4: Pirl3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	92.9	979	2 A35913	regulatory factor
2	102	91.1	592	2 E82759	endo-1,4-beta-gluc
3	100	89.3	207	2 T07381	glycine-rich prote
4	100	89.3	221	2 T04592	glycine-rich cell
5	100	89.3	255	2 B84777	hypothetical prote
6	100	89.3	2783	1 A41948	alpha-fetoprotein
7	98	87.5	266	1 CIRBL	calpain (EC 3.4.22
8	98	87.5	440	2 S71795	transcription fact
9	98	87.5	1969	2 T08875	histidine kinase h
10	97	86.6	151	2 S43296	bone morphogenetic
11	97	86.6	268	1 CTHUL	calpain (EC 3.4.22
12	97	86.6	272	2 T02745	nucleic acid bindi
13	97	86.6	291	1 S31415	glycine-rich prote
14	96	85.7	104	2 J04190	holotricin 3 precu
15	96	85.7	106	2 F84797	hypothetical prote
16	96	85.7	136	2 T29282	hypothetical prote
17	96	85.7	290	2 A01849	hypothetical prote
18	96	85.7	336	1 S18750	chitinase (EC 3.2.
19	96	85.7	443	1 S29334	transcription fact
20	96	85.7	445	1 S31224	transcription fact
21	96	85.7	919	2 A39248	androgen receptor
22	96	85.7	1084	2 T04103	sucrose-phosphate
23	95	84.8	165	1 KNRZG1	glycine-rich cell
24	95	84.8	664	2 J07990	acetylcholinestera
25	94	83.9	162	2 C85356	glycine-rich prote
26	94	83.9	322	2 A45036	single-stranded-DN
27	94	83.9	396	2 T49109	glycine-rich prote
28	94	83.9	431	1 WJHU2G	homeotic protein H
29	94	83.9	433	2 S20963	homeotic protein H

30	94	83.9	1226	2 T24045	hypothetical prote
31	93	83.0	80	2 T10550	hypothetical prote
32	93	83.0	271	2 S34666	glycine-rich prote
33	93	83.0	302	2 C84470	hypothetical prote
34	93	83.0	323	2 S16318	homeotic protein H
35	93	83.0	445	1 A49447	transcription fact
36	93	83.0	901	2 J06093	dead ringer nuclea
37	92	82.1	259	2 T15126	calpain (EC 3.4.22
38	91	81.2	266	1 CIPQL	hypothetical prote
39	91	81.2	273	2 F91083	homeotic protein E
40	91	81.2	333	2 A39065	glycine-rich cell
41	91	81.2	384	1 A26099	female-specific do
42	91	81.2	427	2 A23272	transcription fact
43	91	81.2	495	1 S31223	male-specific doub
44	91	81.2	549	2 B32372	glycine-rich prote
45	90.5	80.8	256	2 T03371	loricrin - mouse
46	90.5	80.8	481	2 A35628	fus-like protein -
47	90.5	80.8	528	2 G02127	zinc finger bindin
48	90.5	80.8	895	2 J07089	nucleic acid bindi
49	90	80.4	273	2 T51145	keratin 1, type II
50	90	80.4	643	1 KRUH2	hypothetical prote
51	90	80.4	877	2 T43449	bulbous pemphigoid
52	90	80.4	1433	2 A46053	homeotic protein s
53	89.5	79.9	284	2 S74256	homeotic protein s
54	89.5	79.9	330	2 S74255	gene M-twist prote
55	89	79.5	206	2 I53066	fibrillarin [valid
56	89	79.5	321	2 A38712	glycin-rich protei
57	89	79.5	408	2 S57483	gene Brr-3b protei
58	89	79.5	410	2 I38502	cell surface glyco
59	89	79.5	779	2 A35006	FUS/CHOP mutant fu
60	88.5	79.0	462	4 S33798	RNA-binding protei
61	88.5	79.0	526	1 S33799	C06G4.3 protein -
62	88	78.6	113	2 S44750	probable glycine-r
63	88	78.6	171	2 H84709	eggshell protein p
64	88	78.6	220	2 A44805	glycine-rich prote
65	88	78.6	252	1 S01821	chitinase (EC 3.2.
66	88	78.6	280	2 A42424	transcription fact
67	88	78.6	420	2 A49642	keratin, 59K type
68	88	78.6	569	1 KRMSE1	keratin, epidermal
69	88	78.6	570	2 S07330	hypothetical prote
70	88	78.6	586	2 T26667	conserved hypothe
71	88	78.6	645	2 AE2855	hypothetical 71.3K
72	88	78.6	672	2 A97437	loricrin - human
73	87.5	78.1	316	1 A38743	hypothetical prote
74	87.5	78.1	341	2 C83578	hypothetical prote
75	87.5	78.1	448	2 T15188	probable single-st
76	87	77.7	139	2 T36594	hypothetical prote
77	87	77.7	301	2 G85928	octamer binding tr
78	87	77.7	420	2 I59234	synGAP-b1 protein
79	87	77.7	1166	2 T13958	Ras-GTPase activat
80	87	77.7	1249	2 T14270	ras GTPase-activat
81	87	77.7	1293	2 T14259	keratin-like prote
82	86.5	77.2	183	2 PNO109	RNA helicase RH11
83	86.5	77.2	224	2 T51742	hypothetical prote
84	86.5	77.2	257	2 C84890	ATP-dependent RNA
85	86.5	77.2	603	2 T45671	glycine-rich RNA-b
86	86	76.8	169	1 S38331	glycine-rich cell
87	86	76.8	338	1 KXNU	F5D14.12 protein -
88	86	76.8	384	2 D86448	hypothetical prote
89	86	76.8	385	2 T20410	homeotic protein U
90	86	76.8	389	2 D26995	hypothetical prote
91	86	76.8	393	2 T20268	hypothetical prote
92	86	76.8	435	2 T15143	hypothetical prote
93	86	76.8	681	2 AB2155	keratin 10, type I
94	85.5	76.3	593	1 KRUH0	glycine-rich RNA-b
95	85	75.9	139	2 S31443	probable RNA-bindi
96	85	75.9	167	2 T05254	glycine-rich RNA-b
97	85	75.9	173	2 S71779	glycine-rich prote
98	85	75.9	173	2 JQ1064	glycine-rich cell
99	85	75.9	183	1 KNRZG2	calpain (EC 3.4.22
100	85	75.9	263	2 A34466	

ALIGNMENTS

```
RESULT 1
A35913
regulatory factor X - human
C;Species: Homo sapiens (man)
C;Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 09-Jul-2004
C;Accession: A35913
R;Reith, W.; Herrero-Sanchez, C.; Kobr, M.; Silacci, P.; Berte, C.; Barras, E.; Fey, S.;
Genes Dev. 4, 1528-1540, 1990
A;Title: MHC class II regulatory factor RFX has a novel DNA-binding domain and a function
A;Reference number: A35913; MUID:91071581; PMID:2253877
A;Accession: A35913
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-979 <REI>
A;Cross-references: UNIPROT:P22670; GB:X58964; NID:g311362; PIDN:CAA41730.1; PID:g33568
C;Keywords: DNA binding; transcription regulation

Query Match      92.9%; Score 104; DB 2; Length 979;
Best Local Similarity 90.0%; Pred. No. 0.012;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
   ||||| ||||| ||||| |||||
DB 381 GGGGGGGGGGGGGGGGGG 400

RESULT 2
E82759
endo-1,4-beta-glucanase XF0818 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82759
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <SIM>
A;Cross-references: UNIPROT:Q9PFP60; GB:AE003921; GB:AE003849; NID:g9105710; PIDN:AAF8362
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Mencia, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
C;Contents: annotation
C;Genetics:
A;Gene: XF0818

Query Match      91.1%; Score 102; DB 2; Length 592;
Best Local Similarity 90.0%; Pred. No. 0.012;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
   ||||| ||||| ||||| |||||
DB 467 GGGGGGGGGGGGGGGGGG 486

RESULT 3
```

T07381

glycine-rich protein Tfm5 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T07381
R;Santino, C.G.; Stanford, G.L.; Conner, T.W.
Plant Mol. Biol. 33, 405-416, 1997
A;Title: Developmental and transgenic analysis of two tomato fruit enhanced genes.
A;Reference number: Z16000; MUID:97201476; PMID:9049262
A;Accession: T07381
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-207 <SAN>
A;Cross-references: UNIPROT:Q43522; EMBL:X95262; NID:g1166449; PIDN:CAA64559.1; PID:g1166
A;Experimental source: cultivar UC82b; fruit
C;Genetics:
A;Gene: Tfm5
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 89.3%; Score 100; DB 2; Length 207;
Best Local Similarity 89.5%; Pred. No. 0.0078; 2; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 19

DB 97 GGGGGGGGGGGGGGGGGG 115

RESULT 4

T04592
glycine-rich cell wall structural protein homolog F23E13.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04592
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Jesse, T.;
submitted to the Protein Sequence Database, March 1998
A;Reference number: Z15378
A;Accession: T04592
A;Molecule type: DNA
A;Residues: 1-221 <BEV>
A;Cross-references: UNIPROT:O65514; EMBL:AL022141
A;Experimental source: cultivar Columbia; BAC clone F23E13
C;Genetics:
A;Map position: 4
A;Note: F23E13.120

Query Match 89.3%; Score 100; DB 2; Length 221;
Best Local Similarity 85.0%; Pred. No. 0.0083;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20

DB 163 GGGGGGGGGGGGGGGGGG 182

RESULT 5

B84777
hypothetical protein A2936120 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84777
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.; N
eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: B84777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-255 <STO>
A;Cross-references: UNIPROT:Q9SIH2; GB:AE002093; NID:g4678224; PIDN:AAD26969.1; GSPDB:GN

A;Gene: At2g36120
A;Map position: 2
C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology

Query Match 89.3%; Score 100; DB 2; Length 255;
Best Local Similarity 89.5%; Pred. No. 0.0093;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
|||||
Db 213 GGGGGGGGGGGGGGGGGG 231

RESULT 6
A41948
alpha-fetoprotein enhancer-binding protein - human
N;Alternate names: ATBF1 protein
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C;Accession: A41948
R;Morinaga, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.
Mol. Cell. Biol. 11, 6041-6049, 1991
A;Title: A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeodomain-like motifs
A;Reference number: A41948; MUID:92049333; PMID:1719379
A;Accession: A41948
A;Molecule type: mRNA
A;Residues: 1-2783 <MOR>
A;Cross-references: GB:D10250; GB:D90395; NID:g219429; PIDN:BA01095.1; PID:g219430
A;Note: sequence extracted from NCBI backbone (NCBIN:66271, NCBIIP:66276)
C;Genetics:
A;Gene: GDB:ATBF1
A;Cross-references: GDB:392090; OMIM:104155
A;Map position: 16q22.3-16q23.1
C;Superfamily: alpha-fetoprotein enhancer-binding protein; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation; zinc finger
F;72-94/Region: zinc finger CCHH motif
F;128-150/Region: zinc finger CCHH motif
F;176-198/Region: zinc finger CCHH motif
F;311-332/Region: zinc finger CCHH motif
F;340-361/Region: zinc finger CCHH motif
F;448-471/Region: zinc finger CCHH motif
F;489-509/Region: zinc finger CCHH motif
F;517-538/Region: zinc finger CCHH motif
F;633-655/Region: zinc finger CCHH motif
F;684-706/Region: zinc finger CCHH motif
F;719-773/Region: zinc finger CCHH motif
F;809-958/Region: glutamine-rich
F;1071-1092/Region: zinc finger CCHH motif
F;1117-1211/Region: proline-rich
F;1232-1288/Domain: homeobox homology <HOX1>
F;1329-1385/Domain: homeobox homology <HOX2>
F;1416-1437/Region: zinc finger CCHH motif
F;1618-1638/Region: zinc finger CCHH motif
F;1728-1784/Domain: homeobox homology <HOX3>
F;1799-1820/Region: zinc finger CCHH motif
F;2033-2089/Domain: homeobox homology <HOX4>
F;2112-2134/Region: zinc finger CCHH motif
F;2545-2566/Region: zinc finger CCHH motif
F;2585-2607/Region: glycine-rich
F;2611-2633/Region: zinc finger CCHH motif
F;2650-2737/Region: serine/threonine-rich

Query Match 89.3%; Score 100; DB 1; Length 2783;
Best Local Similarity 89.5%; Pred. No. 0.063;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
|||||
Db 2587 GGGGGGGGGGGGGGGGGG 2605

RESULT 7
CIRBL

calpain (EC 3.4.22.17) small chain - rabbit
N;Alternate names: calcium-activated neutral proteinase (CANP); calpain light chain; calp
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A24816
R;Emori, Y.; Kawasaki, H.; Imajoh, S.; Kawashima, S.; Suzuki, K.
J. Biol. Chem. 261, 9472-9476, 1986
A;Title: Isolation and sequence analysis of cDNA clones for the small subunit of rabbit c
A;Reference number: A24816; MUID:86250903; PMID:3013892
A;Accession: A24816
A;Molecule type: mRNA
A;Residues: 1-266 <EMO>
A;Cross-references: UNIPROT:P06813; GB:M13364; NID:g164875; PIDN:AAA81565.1; PID:g164876
C;Complex: heterodimer of L (large) and S (small) chains
C;Function:
A;Description: catalyzes the hydrolysis of peptides
A;Note: Cleaves preferentially after tyrosine, methionine, or arginine residues and bef
C;Superfamily: calpain small chain; calmodulin repeat homology
C;Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyd
F;1-54/Domain: glycine-rich <GLY>
F;94-125/Domain: calmodulin repeat homology <EF1>
F;137-169/Domain: calmodulin repeat homology <EF2>
F;170-199/Domain: calmodulin repeat homology <EF3>
F;202-234/Domain: calmodulin repeat homology <EF4>
F;235-266/Domain: calmodulin repeat homology <EF5>

Query Match 87.5%; Score 98; DB 1; Length 266;
Best Local Similarity 80.0%; Pred. No. 0.014;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
|||||
Db 36 GGGGGGGGGGGGGGGGGT 55

RESULT 8
S71795
transcription factor CBF-2 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 16-Aug-2004
C;Accession: S71795
R;Yuasa, J.; Hirano, S.; Yamagata, M.; Noda, M.
Nature 382, 632-635, 1996
A;Title: Visual projection map specified by topographic expression of transcription fact
A;Reference number: S71794; MUID:96338226; PMID:8757134
A;Accession: S71795
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-440 <YUA>
A;Cross-references: UNIPROT:Q98937; EMBL:U47276; NID:g1546783; PIDN:AAB08467.1; PID:g154
C;Superfamily: fork head DNA-binding domain homology
F;143-234/Domain: fork head DNA-binding domain homology <FHD>

Query Match 87.5%; Score 98; DB 2; Length 440;
Best Local Similarity 84.2%; Pred. No. 0.021;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
|||||
Db 113 GGGGAGGGGGGGGGGGG 131

RESULT 9
T08875
histidine kinase homolog DHKB - slime mold (Dictyostelium discoideum)
N;Alternate names: hybrid histidine kinase DHKB
C;Species: Dictyostelium discoideum
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08875
R;Zinda, M.J.; Singleton, C.K.
Dev. Biol. 196, 171-183, 1998
A;Title: The hybrid histidine kinase dhkb regulates spore germination in Dictyostelium d
A;Reference number: 216506; MUID:98248997; PMID:9576830

A;Accession: T08875
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1969 <SIN>
A;Cross-references: UNIPROT:O15763; EMBL:AF024654; NID:g2460282; PID:g2460283
A;Experimental source: strain KAX3
C;Genetics:
A;Gene: dhkB
A;Introns: 790/3
A;Keywords: protein kinase; transmembrane protein
F;1841-1964/Domain: response regulator homology <RRH>

Query Match 87.5%; Score 98; DB 2; Length 1969;
Best Local Similarity 94.4%; Pred. No. 0.071; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;

Qy 2 GGGSGGGGGGGGGGGG 19
Db 549 GGGSGGGGGGGGGGGG 566

RESULT 10
S43296
bone morphogenetic protein-related protein (GDF7) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S43296
R;Storm, E.B.; Huynh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.
Nature 368, 639-643, 1994
A;Title: Limb alterations in brachypodism mice due to mutations in a new member of the T
A;Reference number: S43294; MUID:94195427; PMID:8145850
A;Accession: S43296
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-151 <STO>
A;Cross-references: UNIPROT:P43029; GB:U08339; NID:g488465; PIDN:AAA18780.1; PID:g488466
C;Superfamily: inhibin

Query Match 86.6%; Score 97; DB 2; Length 151;
Best Local Similarity 80.0%; Pred. No. 0.011; Indels 0; Gaps 0;
Matches 16; Conservative 1; Mismatches 3;

Qy 1 GGGSGGGGGGGGGGGG 20
Db 19 GGGSGGGGGGGGGGGG 38

RESULT 11
CIHUL
calpain (EC 3.4.22.17) small chain - human
N;Alternate names: calcium-activated neutral proteinase (CANP)
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A26107; A23650
R;Miyake, S.; Emori, Y.; Suzuki, K.
Nucleic Acids Res. 14, 8805-8817, 1986
A;Title: Gene organization of the small subunit of human calcium-activated neutral prote
A;Reference number: A93648; MUID:87066759; PMID:3024120
A;Accession: A26107
A;Molecule type: DNA
A;Residues: 1-268 <MIY>
A;Cross-references: UNIPROT:P04632; GB:M31502
R;Ohno, S.; Emori, Y.; Suzuki, K.
Nucleic Acids Res. 14, 5559, 1986
A;Title: Nucleotide sequence of a cDNA coding for the small subunit of human calcium-def
A;Reference number: A93631; MUID:86286563; PMID:3016651
A;Accession: A23650
A;Molecule type: mRNA
A;Residues: 1-268 <OHN>
A;Cross-references: EMBL:X04106; NID:g35327; PIDN:CAA27726.1; PID:g35328
C;Genetics:
A;Gene: GDB:CAPN4
A;Cross-references: GDB:119752; OMIM:114170

A;Map position: 19pter-19qter
A;Introns: 70/2; 81/3; 111/3; 131/1; 152/3; 175/3; 202/1; 241/1; 260/3
C;Complex: heterodimer of L (large) and S (small) chains
C;Function:
A;Description: catalyzes the hydrolysis of peptides
A;Note: cleaves preferentially after tyrosine, methionine, or arginine residues and before
C;Superfamily: calpain small chain; calmodulin repeat homology
C;Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hydro
F;1-56/Domain: glycine-rich <GLY>
F;96-127/Domain: calmodulin repeat homology <EF1>
F;139-171/Domain: calmodulin repeat homology <EF2>
F;172-201/Domain: calmodulin repeat homology <EF3>
F;204-236/Domain: calmodulin repeat homology <EF4>
F;237-268/Domain: calmodulin repeat homology <EF5>

Query Match 86.6%; Score 97; DB 1; Length 268;
Best Local Similarity 80.0%; Pred. No. 0.018; Indels 0; Gaps 0;
Matches 16; Conservative 1; Mismatches 3;

Qy 1 GGGSGGGGGGGGGGGG 20
Db 38 GGGSGGGGGGGGGGGG 57

RESULT 12
T02745
nucleic acid binding protein - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02745; T02718
R;Yoon, U.H.; Hahn, J.H.; Yun, C.H.; Eun, M.Y.
submitted to the EMBL Data Library, February 1998
A;Description: The rice genome contains at least two different genes encoding nucleic aci
A;Reference number: Z14712
A;Accession: T02745
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-272 <YOO>
A;Cross-references: UNIPROT:O49228; EMBL:AF047428; NID:g4091116; PID:g4091117
A;Experimental source: strain Ilpoombyeo
R;Yoon, U.H.; Hahn, J.H.; Yun, C.H.; Eun, M.Y.
submitted to the EMBL Data Library, January 1998
A;Description: Cloning and molecular characterization of nucleic acid binding protein ge
A;Reference number: Z14705
A;Accession: T02718
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-181,183-272 <Y02>
A;Cross-references: EMBL:AF045571; NID:g2854124; PID:g2854125
A;Experimental source: strain Ilpoombyeo

Query Match 86.6%; Score 97; DB 2; Length 272;
Best Local Similarity 80.0%; Pred. No. 0.018; Indels 0; Gaps 0;
Matches 16; Conservative 1; Mismatches 3;

Qy 1 GGGSGGGGGGGGGGGG 20
Db 4 GGGSGGGGGGGGGGGG 23

RESULT 13
S31415
glycine-rich protein GRP22 - rape
C;Species: Brassica napus (rape)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S31415
R;Bergeron, D.; Boivin, R.; Baszczynski, C.L.; Bellemare, G.
submitted to the EMBL Data Library, August 1992
A;Description: Characterization and expression of a gene family encoding glycine-rich pr
A;Reference number: S31415
A;Accession: S31415
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-291 <BBR>
A;Cross-references: UNIPROT:Q39337; EMBL:Z15045; NID:gl7820; PIDN:CAA78762.1; PID:gl7821
C;Superfamily: Phaseolus glycinic-rich cell wall protein 1.8

Query Match 86.6%; Score 97; DB 1; Length 291;
Best Local Similarity 84.2%; Pred. No. 0.019;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
|||||:|||||:|||||:|||||
Db 252 GGGGGGGGGGGGGGGGGG 270

RESULT 14
JC4190
holotricin 3 precursor - Holotrichia diomphalia
N;Alternate names: antifungal protein
C;Species: Holotrichia diomphalia
C;Date: 04-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: JC4190
R;Lee, S.Y.; Moon, H.J.; Kurata, S.; Natori, S.; Lee, B.L.
Biol. Pharm. Bull. 18, 1049-1052, 1995
A;Title: Purification and cDNA cloning of an antifungal protein from the hemolymph of H.
A;Reference number: JC4190; MUID:96073722; PMID:8535393
A;Accession: JC4190
A;Molecule type: mRNA
A;Residues: 1-104 <LEB>
A;Cross-references: UNIPROT:Q25055; DBJ:DJ13744; NID:gl088433; PIDN:BA02889.1; PID:d100
C;Comment: This protein is a Gly- and His-rich protein and a constitutive protein of lar
C;Keywords: hemolymph
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-104/Product: holotricin 3 #status predicted <MAT>

Query Match 85.7%; Score 96; DB 2; Length 104;
Best Local Similarity 89.5%; Pred. No. 0.01;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGSGGGGGGGGGGGGGG 20
|||||:|||||:|||||:|||||
Db 60 GGGHGGGGGGGGGGGGGS 78

RESULT 15
F84797
hypothetical protein At2g37830 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84797
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanNaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84797
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-106 <STO>
A;Cross-references: UNIPROT:Q7XJP7; GB:AE002093; NID:g4895201; PIDN:AAD32788.1; GSPDB:GN
C;Genetics:
A;Gene: At2g37830
A;Map position: 2
C;Superfamily: Arabidopsis glycine-rich protein 3

Query Match 85.7%; Score 96; DB 2; Length 106;
Best Local Similarity 84.2%; Pred. No. 0.01;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
|||||:|||||:|||||:|||||
Db 85 GGGGGGGGGGGGGGGGGG 103

RESULT 16
T29282
hypothetical protein C34D4.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29282
R;Du, Z.; Le, T.T.
submitted to the EMBL Data Library, May 1996
A;Description: The sequence of C. elegans cosmid C34D4.
A;Reference number: Z20600
A;Accession: T29282
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-136 <DUZ>
A;Cross-references: UNIPROT:Q18444; EMBL:U58755; PIDN:AAB00696.1; GSPDB:GN00022; CESP:C3-
A;Experimental source: strain Bristol N2; clone C34D4
C;Genetics:
A;Gene: CESP:C34D4.11
A;Map position: 4
A;Introns: 20/1; 66/1; 98/1; 116/1
C;Superfamily: Arabidopsis glycine-rich protein 3

Query Match 85.7%; Score 96; DB 2; Length 136;
Best Local Similarity 84.2%; Pred. No. 0.013;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
|||||:|||||:|||||:|||||
Db 85 GGGGNGGGGGGGGGGGGG 103

RESULT 17
AD1849
hypothetical protein all10341 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD1849
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD1849
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: UNIPROT:Q8VZW4; GB:BA000019; PIDN:BAF72299.1; PID:gl17129686; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all10341

Query Match 85.7%; Score 96; DB 2; Length 290;
Best Local Similarity 84.2%; Pred. No. 0.023;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
|||||:|||||:|||||:|||||
Db 270 GGGGGGGGGGGGGGGGGG 288

RESULT 18
S18750
chitinase (EC 3.2.1.14) precursor - western balsam poplar x cottonwood
C;Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S18750; S17755
R;Davis, J.M.
submitted to the EMBL Data Library, May 1991
A;Reference number: S18750
A;Accession: S18750
A;Molecule type: DNA
A;Residues: 1-336 <DAV>

A;Cross-references: EMBL:X59995
R;Davis, J.M.; Clarke, H.R.G.; Bradshaw Jr., H.D.; Gordon, M.P.
Plant Mol. Biol. 17, 631-639, 1991
A;Title: Populus chitinase genes: structure, organization, and similarity of translated
A;Reference number: S17755; MUID:92003678; PMID:1912489
A;Accession: S17755
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-103 <DAV2>
A;Cross-references: EMBL:X59995
C;Genetics:
A;Introns: 153/1: 203/2
A;Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl
C;Keywords: endoplasmic reticulum; glycosidase; hydrolase; polysaccharide degradation
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-336/Product: chitinase #status predicted <MAT>
F;22-63/Domain: hevein chitin-binding domain homology <HCB>
F;63-82/Region: spacer
F;83-336/Domain: catalytic #status predicted <CAT>
F;88-326/Domain: plant chitinase homology <PCH>

Query Match 85.7%; Score 96; DB 1; Length 336;
Best Local Similarity 84.2%; Pred. No. 0.026;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
||||| ||||| ||||| ||||| |||||
Db 62 GGGGGGGGGGGGGGGGGG 80

RESULT 19
S29334
transcription factor Brn-2 [validated] - human
N;Alternate names: class III POU domain protein brain-2; transcription factor Oct-3
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S29334; S05043; S30296
R;Schreiber, E.; Tobler, A.; Malipiero, U.; Fontana, A.
submitted to the EMBL Data Library, April 1992
A;Description: The human N-Oct 3 cDNA encodes three neuroectodermal cell lineage restrict
A;Reference number: S29334
A;Accession: S29334
A;Molecule type: mRNA
A;Residues: 1-443 <SCH>
A;Cross-references: UNIPROT:P20265; EMBL:Z11933; NID:G35084; PIDN:CAA77990.1; PID:G35085
A;Experimental source: tissue-type brain
R;He, X.; Treacy, M.N.; Simmons, D.M.; Ingraham, H.A.; Swanson, L.W.; Rosenfeld, M.G.
Nature 340, 35-42, 1989
A;Title: Expression of a large family of POU-domain regulatory genes in mammalian brain
A;Reference number: S05042; MUID:89295573; PMID:2739723
A;Accession: S05043
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 280-350:351-404 <HEX>
A;Cross-references: GB:Z11933; NID:G35084
R;Schreiber, E.; Tobler, A.; Malipiero, U.; Schaffner, W.; Fontana, A.
Nucleic Acids Res. 21, 253-258, 1993
A;Title: cDNA cloning of human N-Oct 3, a nervous-system specific POU domain transcript
A;Reference number: S30296; MUID:93181199; PMID:8441633
A;Accession: S30296
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-25,'G',27-443 <SCW>
A;Cross-references: EMBL:Z11933
A;Experimental source: tissue-type brain
C;Genetics:
A;Gene: GDB:POU3P2; OTF7
A;Cross-references: GDB:222816; OMIM:600494
A;Map position: 6q16-6q16
C;Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C;Keywords: alternative initiators; DNA binding; homeobox; nucleus; transcription regula
F;1-443/Product: transcription factor Brn-2 #status experimental <MAT1>

F;68-90/Region: glycine-rich
F;125-149/Region: glutamine-rich
F;151-165/Region: histidine/proline-rich
F;181-443/Product: transcription factor Oct-5a #status experimental <MAT2>
F;200-443/Product: transcription factor Oct-5b #status experimental <MAT3>
F;211-259/Region: histidine/proline-rich
F;269-336/Domain: POU domain homology <POU>
F;355-411/Domain: homeobox homology <HOX>

Query Match 85.7%; Score 96; DB 1; Length 443;
Best Local Similarity 84.2%; Pred. No. 0.032;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
||||| ||||| ||||| ||||| |||||
Db 68 GGGGGGGGGGGGGGGGGG 86

RESULT 20
S31224
transcription factor Brn-2 - mouse
N;Alternate names: class III POU domain protein brain-2
C;Species: Mus musculus (house mouse)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S31224
R;Hara, Y.; Rovescalli, A.C.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A;Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A;Reference number: S31223; MUID:92228768; PMID:1565620
A;Accession: S31224
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-445 <HAR>
A;Cross-references: UNIPROT:P31360; EMBL:M89300; NID:G200446; PIDN:AAA39961.1; PID:G20044
C;Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;68-90/Region: glycine-rich
F;125-151/Region: glutamine-rich
F;153-165/Region: histidine/proline-rich
F;213-261/Region: histidine/proline-rich
F;271-338/Domain: POU domain homology <POU>
F;357-413/Domain: homeobox homology <HOX>

Query Match 85.7%; Score 96; DB 1; Length 445;
Best Local Similarity 84.2%; Pred. No. 0.032;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
||||| ||||| ||||| ||||| |||||
Db 68 GGGGGGGGGGGGGGGGGG 86

RESULT 21
A39248
androgen receptor - human
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1991 #sequence revision 04-Oct-1991 #text_change 09-Jul-2004
C;Accession: A39248; A30328; A40109; A60946; A34942; A27653; A40108; A40494; A32224; A40
R;Lubahn, D.B.; Brown, T.R.; Simental, J.A.; Higgins, H.N.; Migeon, C.J.; Wilson, E.M.; Fr
Proc. Natl. Acad. Sci. U.S.A. 86, 9534-9538, 1989
A;Title: Sequence of the intron/exon junctions of the coding region of the human androgen
A;Reference number: A39248; MUID:90083302; PMID:2594783
A;Accession: A39248
A;Molecule type: DNA
A;Residues: 1-919 <LUB>
A;Cross-references: UNIPROT:P10275; GB:M27423; GB:M27430; NID:G178904; PIDN:AAAS1886.1; I
R;Faber, P.W.; Kuiper, G.G.J.M.; van Rooij, H.C.J.; van der Korput, J.A.G.M.; Brinkmann,
Mol. Cell. Endocrinol. 61, 257-262, 1989
A;Title: The N-terminal domain of the human androgen receptor is encoded by one, large e
A;Reference number: A30328; MUID:89137730; PMID:2917688
A;Accession: A30328
A;Molecule type: DNA
A;Residues: 1-77,79-165,'A',167-389,'L',391-464,473-538 <FAB>

A;Cross-references: GB:M20260
R;Lubahn, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, H.F.; French, F.S.; Wilton, E.M. Science 240, 327-330, 1988
A;Title: Cloning of human androgen receptor complementary DNA and localization to the X chromosome
A;Reference number: A40109; MUID:88178112; PMID:3353727
A;Accession: A40109
A;Molecule type: DNA
A;Residues: 559-624 <LJ2>
A;Cross-references: GB:M20132
R;Kuiper, G.G.J.M.; Faber, P.W.; van Rooij, H.C.J.; van der Korput, J.A.G.M.; Ris-Stalpers, J. Mol. Endocrinol. 2, R1-R4, 1989
A;Title: Structural organization of the human androgen receptor gene.
A;Reference number: A60946; MUID:89322749; PMID:2546571
A;Accession: A60946
A;Molecule type: DNA
A;Residues: 536-540;587-591;626-631;723-726;770-774;814-818;867-870 <KUI>
R;Lubahn, D.B.; Joseph, D.R.; Sar, M.; Tan, J.; Higgs, H.N.; Larson, R.E.; French, F.S. Mol. Endocrinol. 2, 1265-1275, 1988
A;Title: The human androgen receptor: complementary deoxyribonucleic acid cloning, sequence determination, and expression of a cDNA encoding the human androgen receptor
A;Reference number: A34942; MUID:89112208; PMID:3216866
A;Accession: A34942
A;Molecule type: mRNA
A;Residues: 1-919 <LJ3>
A;Cross-references: GB:M20132; NID:g178627; PIDN:AAA51729.1; PID:g178628; GB:J03180
R;Trapman, J.; Klaassen, P.; Kuiper, G.G.J.M.; van der Korput, J.A.G.M.; Faber, P.W.; van Rooij, H.C.J.; Higgs, H.N.; Larson, R.E.; French, F.S. Biochem. Biophys. Res. Commun. 153, 241-248, 1988
A;Title: Cloning, structure and expression of a cDNA encoding the human androgen receptor
A;Reference number: A27653; MUID:88240407; PMID:3377788
A;Accession: A27653
A;Molecule type: mRNA
A;Residues: 468-564, 'K' 566-919 <TRA>
A;Cross-references: GB:M20260; NID:g178891; PIDN:AAA51774.1; PID:g178892
A;Note: the authors translated the codon AAG for residue 565 as Glu
R;Chang, C.; Kokontis, J.; Liao, S. Science 240, 324-326, 1988
A;Title: Molecular cloning of human and rat complementary DNA encoding androgen receptor
A;Reference number: A40108; MUID:88178111; PMID:3353726
A;Accession: A40108
A;Molecule type: mRNA
A;Residues: 557-628 <CHA>
A;Cross-references: GB:M18624
R;Chang, C.; Kokontis, J.; Liao, S. Proc. Natl. Acad. Sci. U.S.A. 85, 7211-7215, 1988
A;Title: Structural analysis of complementary DNA and amino acid sequences of human androgen receptor
A;Reference number: A40494; MUID:89017168; PMID:3174628
A;Accession: A40494
A;Molecule type: mRNA
A;Residues: 1-74, 79-89, 'H' 90-472, 'GGG', 473-474, 'E', 476-644, 'N', 646-919 <CH2>
A;Cross-references: GB:M23263
R;Tilley, W.D.; Marcelli, M.; Wilson, J.D.; McPhaul, M.J. Proc. Natl. Acad. Sci. U.S.A. 86, 327-331, 1989
A;Title: Characterization and expression of a cDNA encoding the human androgen receptor.
A;Reference number: A32224; MUID:89098909; PMID:2911578
A;Accession: A32224
A;Molecule type: mRNA
A;Residues: 1-77, 79-211, 'R' 213-471, 473-919 <TII>
A;Cross-references: GB:M21748; GB:J04150; NID:g178871; PIDN:AAA51771.1; PID:g178872
R;Mowzowicz, I.; Lee, H.J.; Chen, H.T.; Mestayer, C.; Portois, M.C.; Cabrol, S.; Mauval, J. Mol. Endocrinol. 7, 861-869, 1993
A;Title: A point mutation in the second zinc finger of the DNA-binding domain of the androgen receptor.
A;Reference number: A40715; MUID:94019395; PMID:8413310
A;Accession: A40715
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 557-614, 'H' 616-624 <MOW>
A;Cross-references: PIDN:AB28340.1; PID:g425580
C;Genetics:
A;Gene: GDB:AR
A;Cross-references: GDB:120556; OMIM:313700
A;Map position: Xq11-Xq12
A;Intron: 538/2; 589/1; 628/1; 724/1; 772/2; 816/1; 868/3
C;Superfamily: unassigned erba-related proteins; erba transforming protein homology

C;Keywords: DNA binding; steroid binding; transcription regulation; zinc finger
F;557-815/Domain: erba transforming protein homology <ERBA>
F;559-579/Region: zinc finger
F;595-619/Region: zinc finger

Query Match 85.7%; Score 96; DB 2; Length 919;
Best Local Similarity 84.2%; Pred. No. 0.058;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 19
Db 449 GGGGGGGGGGGGGGGGGG 467

RESULT 22

T04103

sucrose-phosphate synthase (EC 2.4.1.14) 1 - rice

C;Species: Oryza sativa (rice)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T04103

R;Sakamoto, M.; Satozawa, T.; Kishimoto, N.; Higo, K.; Shimada, H.; Fujimura, T.

Plant Sci. 112, 207-217, 1995

A;Title: Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS) gene that

A;Reference number: Z15212

A;Accession: T04103

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1084 <SAK>

A;Cross-references: UNIPROT:Q43010; EMBL:D45890; PIDN:BA08304.1

A;Experimental source: subsp. Japonica

C;Genetics:

A;Gene: Spel

A;Map position: 1

A;Introns: 120/3; 200/2; 221/3; 452/3; 492/3; 536/3; 595/3; 613/3; 634/3; 946/3; 989/2

C;Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology

C;Keywords: glycosyltransferase; hexosyltransferase

F;196-680/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 85.7%; Score 96; DB 2; Length 1084;
Best Local Similarity 84.2%; Pred. No. 0.066;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 19
Db 22 GGGGGGGGGGGGGGGGGG 40

RESULT 23

KNR2G1

glycine-rich cell wall structural protein 1 precursor (clone lambda-313) - rice

C;Species: Oryza sativa (rice)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C;Accession: S13385

R;Lei, M.; Wu, R.

Plant Mol. Biol. 16, 187-198, 1991

A;Title: A novel glycine-rich cell wall protein gene in rice.

A;Reference number: S13385; MUID:91370862; PMID:1716496

A;Accession: S13385

A;Molecule type: DNA

A;Residues: 1-165 <LEI>

A;Cross-references: UNIPROT:P25074; EMBL:X53596; NID:g20246; PIDN:CAA37665.1; PID:g20247

C;Genetics:

A;Gene: grp-1

C;Superfamily: glycine-rich cell wall structural protein 1

C;Keywords: cell wall; duplication; structural protein

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-165/Product: glycine-rich cell wall structural protein 1 #status predicted <MAT>

F;30-55/Region: repeat R1

F;56-62/Region: repeat R2

F;62-92/Region: repeat R1

F;93-99/Region: repeat R2

F;100-131/Region: repeat R1

F;132-138/Region: repeat R2

F:139-160/Region: repeat R1

```
Query Match      84.8%; Score 95; DB 1; Length 165;
Best Local Similarity 84.2%; Pred. No. 0.018;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY  2 GGGSGGGSGGGSGGGG 20
    ||||| ||||| ||||| |||||
Db   31 GGGSGGGGGGGGGGGG 49

RESULT 24
JC7990
acetylcholinesterase (EC 3.1.1.7) 1 - green peach aphid
C/Species: Myzus persicae (green peach aphid)
C/Date: 10-Nov-2003 #sequence_revision 10-Nov-2003 #text_change 24-Nov-2003
C/Accession: JC7990
R:Nabeshima, T.; Kozaki, T.; Tomita, T.; Kono, Y.
Biochem. Biophys. Res. Commun. 307, 15-22, 2003
A/Title: An amino acid substitution on the second acetylcholinesterase in the pirimicarb-
A/Reference number: JC7990; PMID:12849975
A/Accession: JC7990
A/Molecule type: mRNA
A/Residues: 1-664 <NAB>
A/Cross-references: GB:AF287291
A/Experimental source: Pirimicarb susceptible strain S14
C/Comment: This enzyme is a target for organophosphate and carbamate insecticides at the
C/Genetics:
A/Gene: MpAChE1
C/Keywords: AChE; acyl pocket; disulphide bond; insecticide resistance

Query Match      84.8%; Score 95; DB 2; Length 664;
Best Local Similarity 84.2%; Pred. No. 0.054;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY  1 GGGSGGGSGGGSGGGG 19
    ||||| ||||| ||||| |||||
Db   33 GGGSGGGGGAGAGSGGGG 51

RESULT 25
C85356
glycine-rich protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: C85356
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: C85356
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-162 <STO>
A/Cross-references: UNIPROT:Q9M0B4; GB:NC_001268; NID:g7269947; PIDN:CAB79764.1; GSPDB:G
C/Genetics:
A/Accession: C85356
A/Map position: 4
C/Superfamily: glycine-rich cell wall structural protein 1

Query Match      83.9%; Score 94; DB 2; Length 162;
Best Local Similarity 84.2%; Pred. No. 0.021;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  1 GGGSGGGSGGGSGGGG 19
    ||||| ||||| ||||| |||||
Db   121 GGGGHHGGGGGGGGGGGG 139

RESULT 26
A45036
single-stranded-DNA-binding protein Pur alpha - human
C/Species: Homo sapiens (man)
```

```
C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C/Accession: A45036
R:Bergemann, A.D.; Ma, Z.W.; Johnson, E.M.
Mol. Cell. Biol. 12, 5673-5682, 1992
A/Title: Sequence of cDNA comprising the human pur gene and sequence-specific single-str
A/Reference number: A45036; MUID:93078769; PMID:1448097
A/Accession: A45036
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-322 <BER>
A/Experimental source: liver
A/Note: sequence extracted from NCBI backbone (NCBIP:119216)
```

```
Query Match      83.9%; Score 94; DB 2; Length 322;
Best Local Similarity 84.2%; Pred. No. 0.037;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  1 GGGSGGGSGGGSGGGG 19
    ||||| ||||| ||||| |||||
Db   29 GSGGGGGGGGGGGGGG 47

RESULT 27
T49109
glycine-rich protein - Arabidopsis thaliana
N/Alternate names: protein AT4g22020
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T49109
R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Ma
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25016
A/Accession: T49109
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-396 <BEV>
A/Cross-references: UNIPROT:O65450; EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g22020
A/Experimental source: cultivar Columbia; BAC clone FlN20
C/Genetics:
A/Gene: ATSP:AT4g22020
A/Map position: 4
C/Superfamily: Phaseolus glycine-rich cell wall protein 1.8
```

```
Query Match      83.9%; Score 94; DB 2; Length 396;
Best Local Similarity 80.0%; Pred. No. 0.044;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY  1 GGGSGGGSGGGSGGGG 20
    ||||| ||||| ||||| |||||
Db   285 GGGGGGGGGGGGGGGGNGS 304
```

```
RESULT 28
WJHU2G
homeotic protein Hox B3 - human
N/Alternate names: homeotic protein Hox 2.7; homeotic protein Hox 2G
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C/Accession: S07543; S15547; D37042
R:Acampora, D.; d'Esposito, M.; Faiella, A.; Pannese, M.; Migliaccio, E.; Morelli, F.; St
Nucleic Acids Res. 17, 10385-10402, 1989
A/Title: The human HOX gene family.
A/Reference number: S07541; MUID:90098876; PMID:2574852
A/Accession: S07543
A/Molecule type: mRNA
A/Residues: 1-431 <ACA>
A/Cross-references: UNIPROT:P14651; EMBL:X16667; NID:g32379; PIDN:CAA34657.1; PID:g32380
R:Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sto
Genome 31, 745-756, 1989
A/Title: Organization of human class I homeobox genes.
A/Reference number: S15036; MUID:90215256; PMID:2576652
A/Accession: S15547
A/Molecule type: DNA
```


A;Residues: 188-253 <BON>
R;Giampolo, A.; Acampora, D.; Zappavigna, V.; Pannese, M.; D'Esposito, M.; Care, A.; Pa
Differentiation 40, 191-197, 1989
A;Title: Differential expression of human HOX-2 genes along the anterior-posterior axis
A;Reference number: A37042; MUID:89378558; PMID:2570724
A;Accession: D37042
A;Molecule type: DNA
A;Residues: 188-253 <GIA>
A;Cross-references: GB:X161175; NID:g32377; PIDN:CAA34297.1; PID:g930068
C;Genetics:
A;Gene: GDB:HOXB3
A;Cross-references: GDB:I20664; OMIM:142966
A;Map position: 17q21.3-17q21.3
A;Introns: 150/1
C;Superfamily: homeotic protein Hox B3; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;152-178/Region: glycine-rich
F;189-245/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 94; DB 1; Length 431;
Best Local Similarity 84.2%; Pred. No. 0.047; 3; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
||||| ||||| ||||| |||||
Db 154 GGGGGGGGGGGGGGGGGG 172

RESULT 29
S20963
homeotic protein Hox B3 - mouse
A;Alternate names: homeotic protein Hox 2.7
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S20963; D42694
R;Sham, M.H.; Hunt, P.; Nonchev, S.; Papalopulu, N.; Graham, A.; Boncinelli, E.; Krumlau
EMBO J. 11, 1825-1836, 1992
A;Title: Analysis of the murine Hox-2.7 gene: conserved alternative transcripts with dif
A;Reference number: S20963; MUID:92258392; PMID:1582411
A;Accession: S20963
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-433 <SHA>
A;Cross-references: UNIPROT:P09026; GB:X66177; GB:S35628; GB:S35738; NID:g312229; PIDN:C
R;Nazarali, A.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 2883-2887, 1992
A;Title: Hox-1.11 and Hox-4.9 homeobox genes.
A;Reference number: A42694; MUID:92212934; PMID:1348361
A;Accession: D42694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 213-238 <NAZ>
A;Note: sequence extracted from NCBI backbone (NCBIN:92310, NCBI:P:92316)
C;Superfamily: homeotic protein Hox B3; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;192-248/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 94; DB 2; Length 433;
Best Local Similarity 84.2%; Pred. No. 0.047;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
||||| ||||| ||||| |||||
Db 156 GGGGGGGGGGGGGGGGGG 174

RESULT 30
T24045
hypothetical protein R08B4.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24045
R;White, S.

submitted to the EMBL Data Library, November 1995

A;Reference number: Z19834
A;Accession: T24045
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1226 <WIL>
A;Cross-references: UNIPROT:Q21835; EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:RO
A;Experimental source: clone R08B4
C;Genetics:
A;Gene: CESP:R08B4.1
A;Map position: X
A;Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 10

Query Match 83.9%; Score 94; DB 2; Length 1226;
Best Local Similarity 84.2%; Pred. No. 0.11;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
||||| ||||| ||||| |||||
Db 854 GGGGGGGGGGGGGGGGGG 872

RESULT 31

T10550

hypothetical protein T12G13.70 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: T10550

R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999

A;Reference number: Z16533

A;Accession: T10550

A;Molecule type: DNA

A;Residues: 1-80 <BEV>

A;Cross-references: UNIPROT:Q9SU7; EMBL:AL080252; GSPDB:GN00062; ATSP:T12G13.70

A;Experimental source: cultivar Columbia; BAC clone T12G13

C;Genetics:

A;Gene: ATSP:T12G13.70

A;Map position: 4

Query Match 83.0%; Score 93; DB 2; Length 80;
Best Local Similarity 88.9%; Pred. No. 0.015;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGGGGGGGGGGGGGGGG 19
||||| ||||| ||||| |||||
Db 18 GGGGGGGGGGGGGGGGGG 35

RESULT 32

S34666

glycine-rich protein - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S34666

R;Brady, K.; Darvill, A.G.; Albersheim, P.

submitted to the EMBL Data Library, July 1993

A;Description: Activation of a tobacco glycine-rich protein gene by a fungal glucan prep

A;Reference number: S34666

A;Accession: S34666

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-271 <BRA>

A;Cross-references: UNIPROT:Q08529; EMBL:X74106; NID:g395146; PIDN:CAA52208.1; PID:g3951

C;Superfamily: Phaseolus glycine-rich protein 1.0

Query Match 83.0%; Score 93; DB 2; Length 271;
Best Local Similarity 84.2%; Pred. No. 0.04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
||||| ||||| ||||| |||||
Db 158 GGGGGGGGGGGGGGGGGG 176

```
RESULT 33
C84470
Hypothetical protein At2g05580 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84470
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84470
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <STO>
A;Cross-references: UNIPROT:Q9SL09; GB:AE002093; NID:g4581166; PIDN:AAD24649.1; GSPDB:GN
C;Genetics:
A;Map position: 2

Query Match 83.0%; Score 93; DB 2; Length 302;
Best Local Similarity 84.2%; Pred. No. 0.043; 3; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
Db 275 GGGGGGGGGGGGGGGG 293

RESULT 34
S16318
homeotic protein Hox 4.6 - mouse
N;Alternate names: homeotic protein Hoxd-11
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Aug-2004
C;Accession: S16318; S40403; S57443
R;Izpisua-Belmonte, J.C.; Falkenstein, H.; Dolle, P.; Renucci, A.; Duboule, D.
EMBO J. 10, 2279-2289, 1991
A;Title: Murine genes related to the Drosophila AbdB homeotic gene are sequentially exp
A;Reference number: S16317; MUID:91293104; PMID:1676674
A;Accession: S16318
A;Molecule type: DNA
A;Residues: 1-323 <12P>
A;Cross-references: UNIPROT:P23813; EMBL:X60761; NID:G51420; PIDN:CAA43173.1; PID:e30097
R;Gerard, M.; Duboule, D.; Zakany, J.
EMBO J. 12, 3539-3550, 1993
A;Title: Structure and activity of regulatory elements involved in the activation of the
A;Reference number: S40403; MUID:94074553; PMID:7902810
A;Accession: S40403
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-323 <GR>
A;Cross-references: EMBL:X71422; NID:g397508; PIDN:CAA50553.1; PID:g397509
R;Takada, S.; Cook, M.; Kranlauf, R.; McMahon, A.P.
submitted to the EMBL Data Library, May 1991
A;Description: Genomic sequence of mouse Hox-4.6.
A;Reference number: S57443
A;Accession: S57443
A;Molecule type: DNA
A;Residues: 'MNDPFCGFSAA',1-323 <TAK>
A;Cross-references: EMBL:X60395; NID:g871427; PIDN:CAA42943.1; PID:g871428
C;Genetics:
A;Gene: Hoxd-11
A;Introns: 246/1
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;252-308/Domain: homeobox homology <HOX>

Query Match 83.0%; Score 93; DB 2; Length 323;
Best Local Similarity 80.0%; Pred. No. 0.046;

RESULT 35
A49447
transcription factor Brn-2 - rat
N;Alternate names: class III POU domain protein brain-2
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 20-Feb-1998
C;Accession: A49447
R;Li, P.; He, X.; Guerrero, M.R.; Mok, M.; Aggarwal, A.; Rosenfeld, M.G.
Genes Dev. 7, 2483-2496, 1993
A;Title: Spacing and orientation of bipartite DNA-binding motifs as potential functional
A;Reference number: A49447; MUID:94102531; PMID:8276233
A;Accession: A49447
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-445 <LI1>
A;Cross-references: GB:L27663; NID:g443687
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:141696)
C;Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;68-90/Region: glycine-rich
F;125-151/Region: glutamine-rich
F;153-165/Region: histidine/proline-rich
F;213-261/Region: histidine/proline-rich
F;271-338/Domain: POU domain homology <POU>
F;357-413/Domain: homeobox homology <HOX>

Query Match 83.0%; Score 93; DB 1; Length 445;
Best Local Similarity 80.0%; Pred. No. 0.059;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 72 GGGGGGGGGGGGGGGG 91

RESULT 36
JC6093
dead ringer nuclear protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: JC6093
R;Gregory, S.L.; Kortschak, R.D.; Kalionis, B.; Saint, R.
Mol. Cell. Biol. 16, 792-799, 1996
A;Title: Characterization of the dead ringer gene identifies a novel, highly conserved f
A;Reference number: JC6093; MUID:96182081; PMID:8622680
A;Accession: JC6093
A;Molecule type: mRNA
A;Residues: 1-901 <GRE>
A;Cross-references: UNIPROT:Q24573; GB:U62542; NID:g1480739; PIDN:AAB05771.1; PID:g148074
A;Experimental source: embryo
C;Genetics:
A;Gene: dri
C;Keywords: DNA binding; embryo

Query Match 83.0%; Score 93; DB 2; Length 901;
Best Local Similarity 84.2%; Pred. No. 0.1;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
Db 203 GTGGGGGGGGGGGGGGGG 221

RESULT 37
T15126
```

hypothetical protein T20B6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15126
R;Beck, C.; Wamsley, P.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid T20B6.
A:Reference number: Z18297
A:Accession: T15126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-259 <BBC>
A:Cross-references: UNIPROT:O02049; EMBL:AF000193; NID:g1946986; PID:g1946988; PIDN:AA05
A:Experimental source: strain Bristol N2; clone T20B6
C:Genetics:
A:Gene: CESP:T20B6.3
A:Map position: 3
A:Introns: 9/2; 231/1
C:Superfamily: Phaseolus glycine-rich protein 1.0

Query Match 82.1%; Score 92; DB 2; Length 259;
Best Local Similarity 84.2%; Pred. No. 0.047;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
Db 110 GGGGGGGGGGGGGGGGGG 128

RESULT 38
CIRP
Calpain (EC 3.4.22.17) small chain - pig
N;Alternate names: calcium-activated neutral proteinase (CANP); calpain light chain; cal
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A25166; B25166
R;Sakihama, T.; Kakidani, H.; Zenita, K.; Yumoto, N.; Kikuchi, T.; Sasaki, T.; Kannagi,
Proc. Natl. Acad. Sci. U.S.A. 82, 6075-6079, 1985
A:Title: A putative Ca2+-binding protein: structure of the light subunit of porcine calp
A:Reference number: A25166; MUID:85298299; PMID:2994060
A:Accession: A25166
A:Molecule type: mRNA
A:Residues: 1-266 <SAK>
A:Cross-references: UNIPROT:P04574; GB:M11778; NID:g164402; PIDN:AAA31010.1; PID:g164403
A:Accession: B25166
A:Molecule type: protein
A:Residues: 2-56;125-143;157-177;247-248;250-256;265-266 <SA2>
C:Complex: heterodimer of L (large) and S (small) chains
C:Function:
A:Description: catalyzes the hydrolysis of peptides
A:Note: cleaves preferentially after tyrosine, methionine, or arginine residues and bef
C:Superfamily: calpain small chain; calmodulin repeat homology
C:Keywords: acetylated amino end; calcium binding; cysteine proteinase; duplication; EF
P.1-54/Domain: glycine-rich <GLY>
P.94-125/Domain: calmodulin repeat homology <EF1>
P.137-169/Domain: calmodulin repeat homology <EF2>
P.170-199/Domain: calmodulin repeat homology <EF3>
P.202-234/Domain: calmodulin repeat homology <EF4>
P.235-266/Domain: calmodulin repeat homology <EF5>
P.1/Modified site: acetylated amino end (Met) #status experimental

Query Match 81.2%; Score 91; DB 1; Length 266;
Best Local Similarity 78.9%; Pred. No. 0.058;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGGGGGGGGGGGGGGGGG 20
Db 37 GGGGGGGGGGGGGGGGGT 55

RESULT 39
P91083
hypothetical membrane protein [imported] - Escherichia coli (strain O157:H7, substrain P

C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F91083
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F91083
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-273 <HAY>
A:Cross-references: UNIPROT:Q8X708; GB:BA000007; PIDN:BA037061.1; PID:g13363109; GSPDB:G
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs3638

Query Match 81.2%; Score 91; DB 2; Length 273;
Best Local Similarity 89.5%; Pred. No. 0.06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGGGGGGGGGGGGGGGG 20
Db 245 GGGGGGGGGGGGGGGGGG 263

RESULT 40
A39065
homeotic protein EVX2 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 17-Oct-1997
C:Accession: A39065
R;D'Esposito, M.; Morelli, F.; Acampora, D.; Migliaccio, E.; Simeone, A.; Boncinelli, E.
Genomics 10, 43-50, 1991
A:Title: EVX2, a human homeobox gene homologous to the even-skipped segmentation gene, is
A:Reference number: A39065; MUID:91257849; PMID:1675198
A:Accession: A39065
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-333 <DAE>
C:Genetics:
A:Gene: GDB:EVX2
A:Cross-references: GDB:127528; OMIM:142991
A:Map position: 2q24.3-2q31
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;46-102/Domain: homeobox homology <HOX>

Query Match 81.2%; Score 91; DB 2; Length 333;
Best Local Similarity 78.9%; Pred. No. 0.07;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
Db 272 GGGGGGGGGGGGGGGGAG 290

Search completed: August 19, 2005, 16:36:12
Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:25:29 ; Search time 174 Seconds

(without alignments)
58.860 Million cell updates/sec

Title: US-09-674-794A-14

Perfect score: 112

Sequence: 1 GGGGGGGGGGGGGGGGGG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trenbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	92.9	979	1	RF1_HUMAN
2	104	92.9	3703	1	ABP1_HUMAN
3	103	92.0	738	2	O02402
4	102	91.1	592	2	Q9PF60
5	100	89.3	118	2	Q9VYS6
6	100	89.3	207	2	Q43522
7	100	89.3	221	2	O65514
8	100	89.3	255	2	Q9SIH2
9	100	89.3	258	2	Q8XLQ8
10	100	89.3	283	2	Q8AVB5
11	100	89.3	288	2	Q7Y1Z0
12	100	89.3	447	2	Q73628
13	100	89.3	541	2	Q87BZ7
14	100	89.3	551	2	Q75HG8
15	100	89.3	1868	2	Q9VWP3
16	99	88.4	318	2	Q38777
17	98	87.5	266	1	C8S1_RABIT
18	98	87.5	287	2	Q6GLA3
19	98	87.5	306	1	RALY_HUMAN
20	98	87.5	355	2	Q7S6V6
21	98	87.5	440	1	FXGA_CHICK
22	98	87.5	465	1	FXD3_MOUSE
23	98	87.5	1969	2	O15763
24	98	87.5	2030	2	Q9VMN8
25	97	86.6	163	2	Q9SUW7
26	97	86.6	165	2	Q9SUX1
27	97	86.6	166	2	Q9SUX0
28	97	86.6	267	2	Q6DI12
29	97	86.6	268	1	C8S1_HUMAN
30	97	86.6	271	2	O49216
31	97	86.6	272	2	O49228

32	97	86.6	272	2	Q7F3M4
33	97	86.6	291	2	Q39337
34	97	86.6	321	2	Q69XV3
35	97	86.6	334	2	Q688M5
36	97	86.6	461	1	GDP7_MOUSE
37	97	86.6	500	1	PO33_HUMAN
38	97	86.6	673	2	Q653C5
39	97	86.6	718	2	Q91TR1
40	96	85.7	61	2	Q8H7A9
41	96	85.7	72	2	Q8ZAB1
42	96	85.7	89	2	Q6ZY22
43	96	85.7	93	2	HOL3_HOLDI
44	96	85.7	104	1	HOL3_HOLDI
45	96	85.7	105	2	Q9GNH2
46	96	85.7	105	2	Q9GP78
47	96	85.7	106	2	Q9GP80
48	96	85.7	106	2	Q7XJP7
49	96	85.7	132	2	Q943G4
50	96	85.7	136	2	Q18444
51	96	85.7	155	2	Q9GND8
52	96	85.7	155	2	Q9GP74
53	96	85.7	156	2	Q9GNB7
54	96	85.7	156	2	Q9GP73
55	96	85.7	157	2	Q9GP77
56	96	85.7	159	2	Q95UM1
57	96	85.7	159	2	Q95UM5
58	96	85.7	161	2	Q95UM6
59	96	85.7	161	2	Q95UX3
60	96	85.7	162	2	Q95UX5
61	96	85.7	163	2	Q95UM6
62	96	85.7	163	2	Q95UM9
63	96	85.7	163	2	Q95UX4
64	96	85.7	164	2	Q95NP2
65	96	85.7	164	2	Q95UM3
66	96	85.7	165	2	Q95NR6
67	96	85.7	165	2	Q95UM2
68	96	85.7	165	2	Q95UX2
69	96	85.7	165	2	Q9GP44
70	96	85.7	167	2	Q9LQZ8
71	96	85.7	168	2	Q95UM8
72	96	85.7	175	2	Q9LSN6
73	96	85.7	192	2	Q92F87
74	96	85.7	220	2	Q6YWF2
75	96	85.7	233	2	Q92NU7
76	96	85.7	239	2	Q69T79
77	96	85.7	281	1	TR2A_MOUSE
78	96	85.7	290	2	Q8YZW4
79	96	85.7	303	1	CHIB_POPTR
80	96	85.7	306	2	Q6Z9T6
81	96	85.7	329	2	Q7Q8W9
82	96	85.7	336	2	Q99N63
83	96	85.7	424	2	Q99N63
84	96	85.7	443	1	PO32_HUMAN
85	96	85.7	445	1	Q86V54
86	96	85.7	531	2	Q9BZG5
87	96	85.7	539	2	Q9NUA2
88	96	85.7	542	2	Q9BZG6
89	96	85.7	544	2	Q9BZG7
90	96	85.7	683	2	Q7QEL7
91	96	85.7	697	2	Q9GRW7
92	96	85.7	698	2	Q9GRX4
93	96	85.7	698	2	Q8QXK8
94	96	85.7	919	1	ANDR_HUMAN
95	96	85.7	1084	2	Q43010
96	96	85.7	1100	2	Q8S064
97	96	85.7	1309	2	Q8PJM4
98	96	85.7	1334	2	Q8PHD6
99	96	85.7	1504	2	Q7PSM9
100	96	85.7			

ALIGNMENTS

```

RESULT 1
ID _RFX1_HUMAN STANDARD; PRT; 979 AA.
AC P22670;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE MHC class II regulatory factor RFX1 (RFX) (Enhancer factor C) (EF-C).
GN Name=RFX1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
DR EMBL; A20498; CAA01506.1; -.
DR EMBL; BC049826; AAH49826.1; -.
DR PIR; A35913; A35913.
DR PDB; 1D27; X-ray; P=438-513.
DR TRANSFAC; T01673; -.
DR Genew; HGNC:9982; RFX1.
DR MIM; 600006; -.
DR GO; GO:0003705; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0006955; P-immune response; TAS.
DR InterPro; IPR007668; RFX1 trans act.
DR InterPro; IPR003150; RFX_DNA_binding.
DR Pfam; PF04589; RFX1 trans act; 1.
DR Pfam; PF02257; RFX_DNA binding; 1.
KW 3D-structure; Activator; DNA-binding; Nuclear protein;
KW Transcription regulation.
KW DOMAIN 381 411 Gly-rich.
FT DNA_BIND 438 528 Experimentally deduced.
FT DOMAIN 920 936 Asp/Glu-rich (acidic).
FT DOMAIN 744 979 Necessary for dimerization.
FT HELIX 439 445
FT STRAND 446 455
FT HELIX 456 469
FT TURN 470 471
FT HELIX 477 487
FT TURN 489 490
FT TURN 492 496
FT TURN 499 500
FT STRAND 503 511
SQ SEQUENCE 979 AA; 104728 MW; 556151F88C6AC9A2 CRC64;
Query Match 92.9%; Score 104; DB 1; Length 979;
Best Local Similarity 90.0%; Pred. No. 0.08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 381 GGGGGGGGGGGGGGGGGG 400
[2]
SEQUENCE FROM N.A.
DR EMBL; X58964; CAA41730.1; -.
DR EMBL; A20498; CAA01506.1; -.
DR PIR; A35913; A35913.
DR PDB; 1D27; X-ray; P=438-513.
DR TRANSFAC; T01673; -.
DR Genew; HGNC:9982; RFX1.
DR MIM; 600006; -.
DR GO; GO:0003705; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0006955; P-immune response; TAS.
DR InterPro; IPR007668; RFX1 trans act.
DR InterPro; IPR003150; RFX_DNA_binding.
DR Pfam; PF04589; RFX1 trans act; 1.
DR Pfam; PF02257; RFX_DNA binding; 1.
KW 3D-structure; Activator; DNA-binding; Nuclear protein;
KW Transcription regulation.
KW DOMAIN 381 411 Gly-rich.
FT DNA_BIND 438 528 Experimentally deduced.
FT DOMAIN 920 936 Asp/Glu-rich (acidic).
FT DOMAIN 744 979 Necessary for dimerization.
FT HELIX 439 445
FT STRAND 446 455
FT HELIX 456 469
FT TURN 470 471
FT HELIX 477 487
FT TURN 489 490
FT TURN 492 496
FT TURN 499 500
FT STRAND 503 511
SQ SEQUENCE 979 AA; 104728 MW; 556151F88C6AC9A2 CRC64;
Query Match 92.9%; Score 104; DB 1; Length 979;
Best Local Similarity 90.0%; Pred. No. 0.08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 381 GGGGGGGGGGGGGGGGGG 400
[3]
IDENTITY BETWEEN RFX1 AND EF-C.
MEDLINE=94019311; PubMed=8413236;
Stegrist C.A., Durand B., Emery P., David E., Hearing P., Mach B.,
Reith W.;
"RFX1 is identical to enhancer factor C and functions as a
transactivator of the hepatitis B virus enhancer.";
Mol. Cell. Biol. 13:6375-6384(1993).
[4]
BINDING TO RPL30 PROMOTER.
MEDLINE=9404774; PubMed=8224874; DOI=10.1016/0378-1119(93)90208-K;
Safrany G., Perry R.P.;
"Transcription factor RFX1 helps control the promoter of the mouse
ribosomal protein-encoding gene rpl30 by binding to its alpha
element.";
Gene 132:279-283(1993).
[5]
SHOWS THAT BLS II IS NOT DUE TO RFX1.
MEDLINE=92375076; PubMed=1508204;
Sanchez-Herrero C., Reith W., Silacci P., Mach B.;
"The DNA-binding defect observed in major histocompatibility complex
class II regulatory mutants concerns only one member of a family of
complexes binding to the X boxes of class II promoters.";
Mol. Cell. Biol. 12:4076-4083(1992).
[6]
FUNCTION: Regulatory factor essential for MHC class II genes
expression. Binds to the X boxes of MHC class II genes. Also binds
to an inverted repeat (ENH1) required for hepatitis B virus genes
```



```

RESULT 3
002402 ID O02402 PRELIMINARY; PRT; 738 AA.
AC O02402;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insoluble protein.
OS Pinctada fucata.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Pterioidea;
OC Pterioidea; Pterioidea; Pinctada.
OX NCBI_TaxID=50426;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97320490; PubMed=9177341; DOI=10.1038/42391;
RA Sudo S., Fujikawa T., Nagakura T., Ohkubo T., Sakaguchi K., Tanaka M.,
RA Nakashima K., Takahashi T.;
RT "Structures of mollusc shell framework proteins.";
RL Nature 387:563-564 (1997).
DR EMBL; D86074; BAA20466.1; --
SQ SEQUENCE 738 AA; 61723 MW; FDF984139BF3BA59 CRC64;

Query Match 92.0%; Score 103; DB 2; Length 738;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGSGGGGS 20
Db 661 GGNGNGGGSGGGSGGGSGGGGS 680

RESULT 4
Q9PF60 ID Q9PF60 PRELIMINARY; PRT; 592 AA.
AC Q9PF60;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Endo-1,4-beta-glucanase.
GN OrderedLocusNames=XF0818;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=945C;
RX Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuranai E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.B.S.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., de Rosa A.J.M.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshahko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

```

```

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159 (2000).
DR EMBL; AB003921; AAF83628.1; --
DR PIR; E82759; E82759.
DR HSSP; P54583; IECE.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001919; Bac_celose-bind.
DR InterPro; IPR008965; Cellul_bind.
DR InterPro; IPR001547; Glyco_hydro_5.
DR Pfam; PF00553; CBM 2; 1.
DR Pfam; PF00150; Cellulase; 1.
DR SMART; SM00637; CBD II; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 592 AA; 59967 MW; 9846DA4EA3B5C89E CRC64;

Query Match 91.1%; Score 102; DB 2; Length 592;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGSGGGGS 20
Db 467 GGGSGGGSGGGSGGGSGGGGS 486

RESULT 5
Q9VYS6 ID Q9VYS6 PRELIMINARY; PRT; 118 AA.
AC Q9VYS6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CG1840-PA (LD12750p).
GN ORFNames=CG1840;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokorva D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evansgelist C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri Y., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,

```


RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Yao Q.A., Ye J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
RT *melanogaster* euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreanek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacieb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003487; AAF48112.1; -;
DR EMBL; AF094783; AAM11136.1; -;
DR FlyBase; FBgn0030351; CG1840.
SQ SEQUENCE 118 AA; 11803 MW; 0568ACA6501716AC CRC64;

Query Match 89.3%; Score 100; DB 2; Length 118;
Best Local Similarity 85.0%; Pred. No. 0.028; Mismatches 0; Gaps 0;
Matches 17; Conservative 0

Oy 1 GGGGGGGGGGGGGGGGGGGG 20
Db 69 GGGGGGGGGGGGGGGGGG 88

RESULT 6

Q43522 Q43522 PRELIMINARY; PRT; 207 AA.
ID Q43522
AC Q43522;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tm5 protein.
DE Names=rfm5;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=fruit;
RC MEDLINE=97201476; PubMed=9049262; DOI=10.1023/A:1005738910743;
RA Santino C.G., Stanford G.L., Conner T.W.;
RT "Developmental and transgenic analysis of two tomato fruit enhanced
RT genes.";
RL Plant Mol. Biol. 33:405-416(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=fruit;
RA Connor T.W.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X95262; CAA64559.1; -;
DR PIR; T07381; T07381.
DR HGSP; P24337; IHYP.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; TRY_alpha_aml; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 207 AA; 18265 MW; A50FE71F8ED4C4D9 CRC64;
Query Match 89.3%; Score 100; DB 2; Length 207;
Best Local Similarity 89.5%; Pred. No. 0.045; Mismatches 0; Gaps 0;
Matches 17; Conservative 0

Oy 1 GGGGGGGGGGGGGGGGGG 19

Db 97 GGGGGGGGGGGGGGGGGG 115

RESULT 7

O65514 O65514 PRELIMINARY; PRT; 221 AA.
ID O65514
AC O65514;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative glycine-rich cell wall protein.
GN Name=P23E13.120; Synonyms=AT4g36230;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoef A.,
RA Hoheisel J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.,
RA Schueller C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoef A.,

```

RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022141; CAAL18129.1; -
DR EMBL; AL161589; CAS80294.1; -
DR PIR; T04592; T04592.
DR InterPro; IPR002952; Eggshell.
DR PRINTS; PR01228; EGGSHLL.
SQ SEQUENCE 221 AA; 20344 MW; DE8DAE2B2C57A9F8 CRC64;

Query Match 89.3%; Score 100; DB 2; Length 221;
Best Local Similarity 85.0%; Pred. No. 0.048;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
Db 163 GGGGGGGGGGGGGGGGGG 182

RESULT 8
Q9SIH2 PRELIMINARY; PRT; 255 AA.
AC Q9SIH2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein At2g36120.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanKen S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Heuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Shinn P., Tang C.C., Torummi M., Wallender E.K., Wong C.,
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
RA Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007135; AAD26969.1; -
DR EMBL; AY136328; AAM96994.1; -
DR EMBL; BT000113; AAN15432.1; -
DR PIR; B84777; B84777.
DR InterPro; IPR002952; Eggshell.
DR PRINTS; PR01228; EGGSHLL.
KW Hypothetical protein.

SQ SEQUENCE 255 AA; 19994 MW; C661A8B1359C6120 CRC64;

Query Match 89.3%; Score 100; DB 2; Length 255;
Best Local Similarity 89.5%; Pred. No. 0.055;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 19
Db 213 GGGGGGGGGGGGGGGGGG 231

RESULT 9
Q8XIQ8 PRELIMINARY; PRT; 258 AA.
AC Q8XIQ8
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein CPE0983.
GN OrderedLocuNames=CPE0983;
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13;
RX MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;
RA Shiba T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003188; BAB80689.1; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 258 AA; 28975 MW; 291ADF32D478BFE9 CRC64;

Query Match 89.3%; Score 100; DB 2; Length 258;
Best Local Similarity 85.0%; Pred. No. 0.055;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
Db 28 GGGGGGGGGGGGGGGGGG 47

RESULT 10
Q8AVB5 PRELIMINARY; PRT; 283 AA.
AC Q8AVB5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sfrs1 protein (Fragment).
GN Name=Sfrs1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

"Sequence and analysis of rice chromosome 4."
RL Nature 420:316-320(2002).
DR EMBL; AB096140; BAC76691.1; -.
DR EMBL; AL606629; CAE01675.2; -.
DR HSSP; P23951; ICNS.
DR Gramene; Q7XT47; -.
DR Gramene; Q7Y120; -.
DR GO; GO:0008061; F.chitin binding; IEA.
DR GO; GO:0004568; F.chitinase activity; IEA.
DR GO; GO:0016998; P.cell wall catabolism; IEA.
DR GO; GO:0006032; P.chitin catabolism; IEA.
DR GO; GO:0009613; P.response to pest, pathogen or parasite; IEA.
DR InterPro; IPR001002; Chitin binding 1.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00187; Chitin_bind_1; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR ProDom; PD000609; Chitin_binding_1; 1.
DR ProDom; PD354900; Glyco_hydro_19; 1.
DR PROSITE; PS00773; CHITINASE_19_1; 1.
DR PROSITE; PS00774; CHITINASE_19_2; 1.
DR PROSITE; PS00026; CHIT_BIND_I_1; 1.
DR PROSITE; PS00941; CHIT_BIND_I_2; 1.
DR Chitin-binding.
KW Chitin-binding.
SQ SEQUENCE 288 AA; 30487 MW; E988A12134910C52 CRC64;

Query Match 89.3%; Score 100; DB 2; Length 288;
Best Local Similarity 85.0%; Pred. No. 0.061; 3; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 3;

QY 1 GGGGGGGGGGGGGGGGGG 20
DB 64 GGGGGGGGGGGGGGGG 83
|||||
|||||
|||||
|||||
|||||

RESULT 12
Q73628 PRELIMINARY; PRT; 447 AA.
AC Q73628;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Brain-2 gene.
GN Name=Brain-2 gene;
OS Anolis carolinensis (Green anole) (American chameleon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Iguanidae; Polychrotidae; Anolis.
ON NCBI_TaxID=28377;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97475689; PubMed=9335144;
RX Nakachi Y., Hayakawa T., Oota H., Sumiyama K., Wang L., Ueda S.;
RT "Nucleotide compositional constraints on genomes generate alanine-,
RT glycine-, and proline-rich structures in transcription factors.";
RL Mol. Biol. Evol. 14:1042-1049(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AB001869; BAA28666.1; -.
DR HSSP; P14859; 1OCT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain like.
DR InterPro; IPR010982; Lambda_like_DNA.
DR InterPro; IPR000327; POU.
DR InterPro; IPR007103; POU homeo.
DR Pfam; PF00046; Homeobox; 1.
DR Pfam; PF00157; POU; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00028; POU DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000583; POU; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.

```

```

DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 447 AA; 47160 MW; AFA362894FCB419 CRC64;

Query Match      89.3%; Score 100; DB 2; Length 447;
Best Local Similarity 85.0%; Pred. No. 0.089;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 69 GGGGGGGGGGGGGGGGGGGG 88

RESULT 13
Q87BZ7
ID Q87BZ7 PRELIMINARY; PRT; 541 AA.
AC Q87BZ7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocName=PD1299;
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RX DOI=10.1128/JB.185.3.1018-1026.2003;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan P.S., Celestino A.V.,
RA da Cunha A.P., Fenille R.C., Ferro J.A., Formighieri E.P., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Saseaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RL J. Bacteriol. 185:1018-1026(2003).
DR EMBL; AE012558; AAC29148.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002372; POQ repeat.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR Pfam; PF01011; POQ; 1.
DR SMART; SM00564; POQ; 7.
KW Complete proteome.
SQ SEQUENCE 541 AA; 55749 MW; 5CB0BB559803426F CRC64;

Query Match      89.3%; Score 100; DB 2; Length 541;
Best Local Similarity 85.0%; Pred. No. 0.11;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
Db 455 GGGGGGGGGGGGGGGGGGGG 474

RESULT 14
Q75HG8
ID Q75HG8 PRELIMINARY; PRT; 551 AA.
AC Q75HG8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative prohibitin.
GN Name=OSNBA0024F18.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135594; AAR89849.1; -.
DR InterPro; IPR001107; Band 7.
DR Pfam; PF01145; Band 7; 1.
DR SMART; SM02444; PHB; 1.
SQ SEQUENCE 551 AA; 59192 MW; 34C1EC87BB3CA297 CRC64;

Query Match      89.3%; Score 100; DB 2; Length 551;
Best Local Similarity 89.5%; Pred. No. 0.11;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 19
Db 42 GGGGGGGGGGGGGGGGGGGG 60

RESULT 15
Q9VWP3
ID Q9VWP3 PRELIMINARY; PRT; 1868 AA.
AC Q9VWP3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG7282-FA.
GN ORFNames=CG7282;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherzer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,

```

RA Jalali M., Kaluah P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levcintosh A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.A., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheier F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003510; AAP48895.2; -;
DR FlyBase; FBgn0030967; CG7282.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003682; F:chromatin binding; IEA.
DR GO; GO:0006333; P:chromatin assembly or disassembly; IEA.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR001412; tRNA-synt_1.
DR SMART; SM00298; CHROMO; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; UNKNOWN 1.
SQ SEQUENCE 1868 AA; 193058 MW; 9E1F9E13C9F7F97A CRC64;

Query Match 89.3%; Score 100; DB 2; Length 1868;
Best Local Similarity 85.0%; Pred. No. 0.31;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGGGG 20
DB 158 GGGGGGGGGGGGGGGGGGGGGG 177

RESULT 16
Q38777 PRELIMINARY; PRT; 318 AA.
AC Q38777;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chitinase precursor (fragment).
GN Name=chitinase;
OS Allium sativum (Garlic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
OC Allium.
OX NCBI_TaxID=4682;
RN [1]
RP SEQUENCE FROM N.A.
RA van Damme E.J.M., Willems P., Peumans W.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M94105; AAA32641.1; -;
DR HGSP; P23951; 2BAA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0004568; F:chitinase activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
DR GO; GO:0009613; P:response to pest, pathogen or parasite; IEA.
DR InterPro; IPR001002; Chitin_binding_1.
DR Pfam; PF00187; Chitin_bind_1; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR PRINTS; PR00451; CHITINBINDING.
DR PRODOM; PD000609; Chitin_binding_1; 1.
DR PRODOM; PD354900; Glyco_hydro_19; 1.
DR SMART; SM00270; ChtBD1; 1.
DR PROSITE; PS00773; CHITINASE_19_1; 1.
DR PROSITE; PS00774; CHITINASE_19_2; 1.
DR PROSITE; PS00026; CHIT_BIND_1; 1.
DR PROSITE; PS00941; CHIT_BIND_2; 1.
KW Chitin-binding; Signal.
FT SIGNAL 1 17
FT CHAIN 18 >318 chitinase.
FT NON_TER 318 318
SQ SEQUENCE 318 AA; 33175 MW; 097C3DBD772468D4 CRC64;

Query Match 88.4%; Score 99; DB 2; Length 318;
Best Local Similarity 85.0%; Pred. No. 0.081;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGGGG 20
DB 57 GGGGGGGGGGGGGGGGGGGGGG 76

RESULT 17
CSSI_RABIT STANDARD; PRT; 266 AA.
ID CSSI_RABIT
AC P06813;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Calpain small subunit 1 (CSSI) (Calcium-dependent protease small
subunit 1) (Calcium-dependent protease small subunit) (CDPS) (Calpain
regulatory subunit) (Calcium-activated neutral proteinase small
subunit) (CAMP small subunit).
DE Name=CAPNS1; Synonyms=CAPN4;
GN


```

[2]
SEQUENCE FROM N.A. (ISOFORM 1).
Vaughan J.H.;
Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;
RA Deloukas P., Matthews L.H., Aheurst J.L., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leinvaelt M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RA "The DNA sequence and comparative analysis of human chromosome 20.";
RA Nature 414:865-871(2001).
[4]
SEQUENCE OF 85-306 FROM N.A. (ISOFORM 1), AND AUTOIMMUNE DISEASE.
TISSUE=Lymphocytes;
RX MEDLINE=98018738; PubMed=9376072; DOI=10.1006/jaut.1997.9996;
RA Rhodes G.H., Valbracht J.R., Nguyen M.-D., Vaughan J.H.;
RA "The p542 gene encodes an autoantigen that cross-reacts with EBNA-1 of
RT the Epstein Barr virus and which may be a heterogeneous nuclear
RT ribonucleoprotein.";
RA J. Autoimmun. 10:447-454(1997).
[5]
SEQUENCE OF 227-253, AND DETERMINATION OF AUTOANTIGENIC EPITOPE.
RX MEDLINE=95190029; PubMed=7533788;
RA Vaughan J.H., Valbracht J.R., Nguyen M.-D., Handley H.H., Smith R.S.,
RA Patrick K., Rhodes G.H.;
RT "Epstein-Barr virus-induced autoimmune responses. I. Immunoglobulin M
RT autoantibodies to proteins mimicking and not mimicking Epstein-Barr
RT virus nuclear antigen-1.";
RT J. Clin. Invest. 95:1306-1315(1995).
CC -!- FUNCTION: Probable-RNA binding protein. Could be a heterogeneous
CC nuclear ribonucleoprotein (hnRNP).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=2;
CC IsoId=Q9UKM9-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UKM9-2; Sequence=VSP_005804;
CC -!- TISSUE SPECIFICITY: Expressed in heart, brain, lung, liver,
CC skeletal muscle, kidney and pancreas. Weakly expressed in
CC placenta.
CC -!- DISEASE: Autoantigen found in infectious mononucleosis caused by
CC Epstein-Barr virus. An epitope recognized by B-cells, which cross-
CC react with the BKRFL protein (EBNA-1 nuclear protein) of Epstein-
CC Barr virus has been identified.
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -!- CAUTION: Ref.4 (CAC29371) sequence differs from that shown due to
CC erroneous gene model prediction.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; AABX01000321; EAA31335.1; -.
DR HSSP; Q9S7B3; 1EHD.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001211; PhospholipaseA2.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS00118; PA2_HIS; UNKNOWN_1.
SQ SEQUENCE 355 AA; 33593 MW; CDF5AB4DFDF6A7EC CRC64;

Query Match 87.5%; Score 98; DB 2; Length 355;
Best Local Similarity 85.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;
Matches 17; Conservative 0

Oy 1 GGGGSGGGSGGGSGGGSGGS 20
||||| ||||| |||||
Db 139 GGGGSGGGSGGGSGGGSGGS 158

RESULT 21
FXGA_CHICK STANDARD; PRT; 440 AA.
ID FXGA_CHICK AC Q98937;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Forkhead box protein GLA (forkhead-related protein FKHL2)
DE (Transcription factor BF-2) (Brain factor 2) (BF2) (CBF-2) (T-14-6).
GN Names=FOXGLA; Synonyms=FKHL2, HFHB2;
OS Gallus gallus (Chicken).
OC Chordata; Craniata; Vertebrata; Euteleostomi;
OC Arcohsauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Retina;
RX MEDLINE=96338226; PubMed=8757134; DOI=10.1038/382632a0;
RA Yuasa J., Hirano S., Yamagata M., Noda M.;
RT "Visual projection map specified by topographic expression of transcription factors in the retina."
RL Nature 382:632-635(1996).
CC -!- FUNCTION: May determine the nasotemporal axis of the retina, and consequently specify the topographical projection of the retinal ganglion-cell axons to the tectum by controlling expression of their target genes.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Retina.
CC -!- DEVELOPMENTAL STAGE: Can be detected in regions including primordial retina and neuroepithelium by embryonic day 2 (E2). At E3, expressed in the temporal retina and associated pigment epithelium as well as in part of the diencephalon, and at E7 is expressed in retinal ganglion cells. Levels begin to decline from E4 and almost disappear by E10.
CC -!- SIMILARITY: Contains 1 fork-head domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U47276; AAB08467.1; -.
PIR; S71795; S71795.
HSSP; Q63245; 2HDC.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement or send an email to license@isb-sib.ch).

EMBL; U41047; AAA87569.1; -.
EMBL; AF067421; AAC28352.1; -.
HSSP; Q63245; 2HDC.
TRANSFAC; T04166; ..
MGD; MGI:1347473; Foxd3.
InterPro; IPR001766; TF_Fork_head.
InterPro; IPR009058; Wing_hlx_DNA_bnd.
PFam; PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
ProDom; PD000425; TF_Fork_head; 1.
SMART; SM00339; FH; 1.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00657; FORK_HEAD_3; 1.
PROSITE; PS00657


```

FT DOMAIN 106 115 Poly-Gly.
FT DNA BIND 131 225 Fork-head.
FT DOMAIN 252 257 Poly-Ala.
FT DOMAIN 265 270 Poly-Ala.
FT DOMAIN 275 281 Poly-Ala.
FT DOMAIN 380 399 Poly-Gly.
FT DOMAIN 447 457 Poly-Ala.
SQ SEQUENCE 465 AA; 47092 MW; 6F8B5B3D8C7564D CRC64;

Query Match 87.5%; Score 98; DB 1; Length 465;
Best Local Similarity 89.5%; Pred. No. 0.14; 2; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGG 19
|||||
Db 380 GGGGGGGGGGGGGGGGG 398

RESULT 23
O15763 PRELIMINARY; PRT; 1969 AA.
ID O15763
AC O15763;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hybrid histidine kinase DHKB.
GN Name=dhkb;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAX3;
RX MEDLINE=98248997; PubMed=9576830; DOI=10.1006/dbio.1998.8854;
RA Zinda M.J., Singleton C.K.;
RT "The hybrid histidine kinase dhkB regulates spore germination in Dictyostelium discoideum."
RL Dev. Biol. 196:171-183 (1998).
CC -i- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; AF024654; AAB71889.1; -.
DR PIR; T08875; T08875.
DR HSSP; P39928; 10XK.
DR DictyBase; DB0215358; dhkb.
DR GO; GO:0016020; C:membrane. IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HsKA; 1.
DR Pfam; PF00072; Response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HsKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS0109; HIS_KIN; 1.
DR PROSITE; PS0110; RESPONSE_REGULATOR; 1.
KW Kinase; Phosphorylation; Sensory transduction.
SQ SEQUENCE 1969 AA; 219024 MW; 8E7A7952AB1BB52B CRC64;

Query Match 87.5%; Score 98; DB 2; Length 1969;
Best Local Similarity 94.4%; Pred. No. 0.49;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGGGGGGGGGGGGGGGG 19
|||||
Db 549 GGGGGGGGGGGGGGGGG 566

RESULT 24
Q9VWN8 PRELIMINARY; PRT; 2030 AA.
ID Q9VWN8
AC Q9VWN8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG32542-PA.
GN ORFNames=CG32542;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.L., Abrial J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R., Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence."

```

```
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RX SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RX SEQUENCE FROM N.A.
RP MEDLINE=22426069; PubMed=12537572;
RA Mera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RX SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RX SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003510; AAF48900.2; -.
DR FlyBase; FBgn0052542; CG32542.
DR InterPro; IPR007527; Znf_SWIM.
DR Pfam; PF04434; SWIM; 1.
SQ SEQUENCE 2030 AA; 216704 MW; D5D2F3CCC8FA2241 CRC64;

Query Match 87.5%; Score 98; DB 2; Length 2030;
Best Local Similarity 84.2%; Pred. No. 0.5;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 19
DB 1282 GGGGNGGGGGGGGGGGGG 1300

RESULT 25
Q95UW7 PRELIMINARY; PRT; 163 AA.
AC Q95UW7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE No on or off transient A (Fragment).
GN Name=nonA;
OS Drosophila americana americana.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=95109;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NN97.7;
RA Huttunen S., Vieira J., Hoikkala A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY012598; AAG48868.1; -.
DR FlyBase; FBgn0046723; Dtex\nonA.
FT NON_TER 163 163
FT NON_TER 163 163
SQ SEQUENCE 163 AA; 15892 MW; BF6F99C45E412EB5 CRC64;

Query Match 86.6%; Score 97; DB 2; Length 163;
Best Local Similarity 80.0%; Pred. No. 0.067;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
DB 71 GGGGGGGGGGGGGGGGGGGA 90

RESULT 26
Q95UX1 PRELIMINARY; PRT; 165 AA.
AC Q95UX1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE No on or off transient A (Fragment).
GN Name=nonA;
OS Drosophila americana americana.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=95109;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=0951.0;
RA Huttunen S., Vieira J., Hoikkala A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY012594; AAG48864.1; -.
DR FlyBase; FBgn0046746; Dame\nonA.
FT NON_TER 165 165
FT NON_TER 165 165
SQ SEQUENCE 165 AA; 16006 MW; E2C2B5D2ACD193CA CRC64;

Query Match 86.6%; Score 97; DB 2; Length 165;
Best Local Similarity 80.0%; Pred. No. 0.068;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
DB 73 GGGGGGGGGGGGGGGGGGGA 92

RESULT 27
Q95UX0 PRELIMINARY; PRT; 166 AA.
AC Q95UX0
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE No on or off transient A (Fragment).
GN Name=nonA;
OS Drosophila novamexicana.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=47314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1031.0;
RA Huttunen S., Vieira J., Hoikkala A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY012595; AAG48865.1; -.
DR FlyBase; FBgn0046734; Dnov\nonA.
FT NON_TER 166 166
FT NON_TER 166 166
SQ SEQUENCE 166 AA; 16148 MW; F484F3F17A192B28 CRC64;

Query Match 86.6%; Score 97; DB 2; Length 166;
Best Local Similarity 80.0%; Pred. No. 0.068;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 20
DB 74 GGGGGGGGGGGGGGGGGGGA 93
```


[6]
X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
MEDLINE=2010516; PubMed=10639123; DOI=10.1073/pnas.97.2.588;
Strobl S., Fernandez-Catalan C., Braun M., Huber R., Masumoto H.,
Nakagawa K., Irie A., Sorimachi H., Bourenkow G., Bartunik H.,
Suzuki K., Bode W.;
"The crystal structure of calcium-free human m-calpain suggests an
electrostatic switch mechanism for activation by calcium.";
Proc. Natl. Acad. Sci. U.S.A. 97:588-592(2000).
-1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
catalyze limited proteolysis of substrates involved in
cytoskeletal remodelling and signal transduction.
-1- SUBUNIT: Heterodimer of a large (catalytic) and a small
(regulatory) subunit.
-1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
membrane upon calcium binding (By similarity).
-1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
-1- DATABASE: NAME=CaBP; NOTE=Calpain;
WWW="http://structbio.vanderbilt.edu/cabp_database/general/prot_pages/calpain.h
-1- DATABASE: NAME=Calpains homepage;
WWW="http://ag.arizona.edu/calpains".

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X04106; CAA27726.1; -
EMBL; M31511; AAA35646.1; -
EMBL; M31502; AAA35646.1; JOINED.
EMBL; M31503; AAA35646.1; JOINED.
EMBL; M31504; AAA35646.1; JOINED.
EMBL; M31505; AAA35646.1; JOINED.
EMBL; M31506; AAA35646.1; JOINED.
EMBL; M31507; AAA35646.1; JOINED.
EMBL; M31508; AAA35646.1; JOINED.
EMBL; M31509; AAA35646.1; JOINED.
EMBL; M31510; AAA35646.1; JOINED.
EMBL; AD001527; AAB51183.1; -
EMBL; BT009775; AAP8777.1; -
EMBL; AC002984; AAB81546.1; -
EMBL; BC000592; AAH00592.1; -
EMBL; BC007779; AAH07779.1; -
EMBL; BC011903; AAH11903.1; -
EMBL; BC017308; AAH17308.1; -
EMBL; BC018931; AAH18931.1; -
EMBL; BC021933; AAH21933.1; -
EMBL; BC023643; AAH23643.1; -
EMBL; BC064998; AAH64998.1; -
PIR; A26107; CIHUL.
PDB; 1KFU; X-ray; S=85-268.
PDB; 1KEX; X-ray; S=85-268.
SWISS-2DPAGE; P04632; HUMAN.
OGP; P04632; -
Genew; HGNC:1481; CAPNS1.
H-InvDB; HIX0015056; -
MIM; 114170; -
GO; GO:0004198; F:calpain activity; TAS.
GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
InterPro; IPR002048; EF-hand.
InterPro; IPR010983; EF-hand_like.
Pfam; PF00036; ehand; 3.
ProDom; PD000012; EF-hand; 1.
SMART; SM00054; EFh; 3.
PROSITE; PS00018; EF_HAND; 2.
3D-structure; Calcium-binding; Repeat.
DOMAIN 1 66 Gly-rich (hydrophobic).
CA BIND 152 163 EF-hand 1.
CA BIND 182 193 EF-hand 2.
DOMAIN 217 228 Ancestral calcium site 3 (Potential).
FT

FT DOMAIN 247 260 Ancestral calcium site 4 (Potential).
FT DOMAIN 10 26 Poly-Gly.
FT DOMAIN 35 56 Poly-Gly.
FT DOMAIN 78 83 Poly-Pro.
FT CONFLICT 5 5 N -> D (in Ref. 5; AAH64998).
FT CONFLICT 27 27 N -> G (in Ref. 5; AAH64998).
FT CONFLICT 34 34 S -> G (in Ref. 5; AAH64998).
FT CONFLICT 261 268 WLQLTWYS -> VRTPLGVGCLGGPHPSALHTSSLSQSPS
SYFASRPWYRAKGLVLLGFPVLTLPPLPSGCS (in
Ref. 5; AAH1903).
Y -> F (in Ref. 5; AAH21933).
FT CONFLICT 267 267
FT HELIX 96 108
FT TURN 109 111
FT STRAND 115 115
FT HELIX 117 124
FT TURN 127 129
FT TURN 132 133
FT HELIX 141 151
FT STRAND 159 159
FT HELIX 162 180
FT TURN 183 184
FT STRAND 189 189
FT HELIX 191 193
FT HELIX 194 200
FT TURN 201 202
FT HELIX 207 217
FT STRAND 219 219
FT TURN 220 222
FT STRAND 223 224
FT HELIX 226 246
FT STRAND 253 256
FT HELIX 258 266
SQ SEQUENCE 268 AA; 28316 MW; 17B87A8E47A90632 CRC64;
Query Match 86.6%; Score 97; DB 1; Length 268;
Best Local Similarity 80.0%; Pred. No. 0.1;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGGSGGGSGGGSGGGGS 20
|||||
DB 38 GGGSGGGSGGGSGGGGT 57
RESULT 30
O49216 PRELIMINARY; PRT; 271 AA.
ID O49216
AC O49216;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nucleic acid binding protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoon U.H., Hahn J.H., Yun C.-H., Eun M.Y.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF045571; AAC98962.1; -
DR Gramene; O49216; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011011; FYVE PHD_Znf.
DR InterPro; IPR001965; Znf_PHD.
DR Pfam; PF00628; PHD; 1.
DR SMART; SM00249; PHD; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS0016; ZF_PHD_2; 1.
SQ SEQUENCE 271 AA; 29476 MW; 0447690D466F1A32 CRC64;
Query Match 86.6%; Score 97; DB 2; Length 271;

```

Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGG 20
Db 4 GGGGGGGGGGGGGGGGGGGG 23

RESULT 31
ID O49228 PRELIMINARY; PRT; 272 AA.
AC O49228;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nucleic acid binding protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
SQ SEQUENCE 272 AA; 29604 MW; 964DFC13AFC7E2 CRC64;
[1]
SEQUENCE FROM N.A.
RP Yoon U.H., Hahn J.H., Yun C.-H., Eun M.Y.;
RA Submitted (DSC-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF047428; AAC98969.1; -.
DR PIR; T02745; T02745.
DR Gramene; O49228; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011011; FYVE_PHD_Znf.
DR InterPro; IPR001965; Znf_PHD.
DR Pfam; PF00628; PHD; 1.
DR SMART; SM00249; PHD; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS00016; ZF_PHD_2; 1.
SQ SEQUENCE 272 AA; 29604 MW; 964DFC13AFC7E2 CRC64;

Query Match 86.6%; Score 97; DB 2; Length 272;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGG 20
Db 4 GGGGGGGGGGGGGGGGGGGG 23

RESULT 32
ID Q7F3M4 PRELIMINARY; PRT; 272 AA.
AC Q7F3M4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Nucleic acid binding protein.
GN Name=P0434C04.21; Synonyms=R1099D03.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
SQ SEQUENCE FROM N.A.
RP PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arica K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh Y., Iwabuuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,

```

```

RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yanane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RL "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003346; BAC0644.1; -.
DR EMBL; AP003431; BAB92630.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011011; FYVE_PHD_Znf.
DR InterPro; IPR001965; Znf_PHD.
DR Pfam; PF00628; PHD; 1.
DR SMART; SM00249; PHD; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS00016; ZF_PHD_2; 1.
SQ SEQUENCE 272 AA; 29604 MW; 964DFC13AFC7E2 CRC64;

Query Match 86.6%; Score 97; DB 2; Length 272;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGG 20
Db 4 GGGGGGGGGGGGGGGGGGGG 23

RESULT 33
ID Q39337 PRELIMINARY; PRT; 291 AA.
AC Q39337;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycine-rich protein_(Aal-291) precursor.
GN Name=GRP22;
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Westar;
RA Bergeron D., Boivin R., Baszczynski C.L., Bellemare G.;
RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z15045; CAA78762.1; -.
DR PIR; S31415; S31415.
DR InterPro; IPR002952; Eggshell.
DR PRINTS; PR01228; EGGSHLL.
KW Signal.
FT SIGNAL
SQ SEQUENCE 291 AA; 22659 MW; 305C7ECE0DFEA264 CRC64;

Query Match 86.6%; Score 97; DB 2; Length 291;
Best Local Similarity 84.2%; Pred. No. 0.11;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGG 19
Db 252 GGGGGGGGGAGGGGGGGGG 270

RESULT 34
ID Q69XV3 PRELIMINARY; PRT; 321 AA.
AC Q69XV3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative glycine-rich cell wall structural protein.
GN Name=P0040H10.37;

```

```

OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
  clone:P0040H10.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003539; BAD35374.1; -;
SQ SEQUENCE 321 AA; 28605 MW; DD83382025B2F649 CRC64;

Query Match      86.6%; Score 97; DB 2; Length 321;
Best Local Similarity 80.0%; Pred. No. 0.12;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 129 GGGGGGGGGGGGGGGG 148

RESULT 35
Q688M5 PRELIMINARY; PRT; 334 AA.
AC Q688M5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Putative chitinase.
GN Name=P0605G01.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.-C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hou C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RT "Oryza sativa PAC P0605G01 genomic sequence.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC132492; AAU10806.1; -;
DR InterPro; IPR001002; Chitin_binding_1.
DR Pfam; PF00187; Chitin_bind_1; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR PRINTS; PR00451; CHITINBINDING.
DR ProDom; PD000609; Chitin_binding_1; 1.
DR ProDom; PD354900; Glyco_hydro_19; 1.
DR SMART; SM00270; ChcBD1_1.
DR PROSITE; PS00774; CHITINASE_19_2; 1.
DR PROSITE; PS00026; CHIT_BIND_I_1; 1.
KW Chitin-binding.
SQ SEQUENCE 334 AA; 34401 MW; B147004C31999D14 CRC64;

Query Match      86.6%; Score 97; DB 2; Length 334;
Best Local Similarity 80.0%; Pred. No. 0.13;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGG 20
Db 66 GGGGGGGGGGGGGGGG 85

RESULT 36
GDF7_MOUSE
ID _GDF7_MOUSE STANDARD; PRT; 461 AA.

```

```

AC P43029; Q7TNX4; Q99MY1;
DT 01-NOV-1995 (Rel. 32, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Growth/differentiation factor 7 precursor (GDF-7).
GN Name=Gdf7; Synonyms=Gdf-7;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=BALB/c; TISSUE=Testis;
RA Guo J.H., Yu L.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-441 FROM N.A. (ISOFORM 1).
RC STRAIN=TT2;
RX MEDLINE=21136583; PubMed=11238730;
RA Watakabe A., Fujita H., Hayashi M., Yamamori T.;
RT "Growth/differentiation factor 7 is preferentially expressed in the
  primary motor area of the monkey neocortex.";
RL J. Neurochem. 76:1455-1464 (2001).
RN [3]
RP SEQUENCE OF 311-461 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=94195427; PubMed=8145850; DOI=10.1038/368639a0;
RA Storm E.E., Huynh T.V., Copeland N.G., Jenkins N.A., Kingsley D.M.,
RA Lee S.-J.;
RT "Limb alterations in brachypodism mice due to mutations in a new
  member of the TGF beta-superfamily.";
RL Nature 368:639-643 (1994).
CC -1- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P43029-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P43029-2; Sequence=VSP_010764;
CC -1- SIMILARITY: Belongs to the TGF-beta family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF525752; AAP97721.1; -.
DR EMBL; AF254571; AAK30843.1; -.
DR EMBL; AF254570; AAK30843.1; JOINED.
DR EMBL; U08339; AAA18780.1; -.
DR PIR; S43296; S43296.
DR HSSP; P12643; 3BMP.
DR MGD; MGI:95690; Gdf7.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007411; P:axon guidance; IDA.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR Pfam; PF00019; TGF_beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Alternative splicing; Cytokine; Glycoprotein; Growth factor; Signal.
FT SIGNAL 1 19 Potential.
FT PROPEP 20 315 Potential.
FT CHAIN 316 461 Growth/differentiation factor 7.
FT DISULFID 360 426 By similarity.
FT DISULFID 389 458 By similarity.

```

DR	InterPro; IPR007103; POU homeo.
DR	Pfam; PF00046; Homeobox_1.
DR	Pfam; PF00157; Pou; 1.
DR	PRINTS; PR00028; POU DOMAIN.
DR	ProDom; PD000010; Homeobox; 1.
DR	ProDom; PD000583; POU; 1.
DR	SMART; SM00389; Hox; 1.
DR	SMART; SM00352; POU; 1.
DR	PROSITE; PS00027; HOMEBOX_1; 1.
DR	PROSITE; PS50071; HOMEBOX_2; 1.
DR	PROSITE; PS00035; POU 1; 1.
DR	PROSITE; PS00465; POU 2; 1.
KW	DNA-binding; Homeobox; Nuclear protein.
FT	DOMAIN 28 50 Poly-Gly.
FT	DOMAIN 104 115 Poly-Ala.
FT	DOMAIN 136 144 Poly-Pro.
FT	DOMAIN 174 179 Poly-Pro.
FT	DOMAIN 189 204 Poly-Ala.
FT	DOMAIN 239 250 Poly-Gly.
FT	DOMAIN 270 294 His-rich.
FT	DOMAIN 270 281 Poly-His.
FT	DOMAIN 295 304 Poly-Gly.
FT	DOMAIN 318 388 POU.
FT	DNA_BIND 406 465 Homeobox.
FT	CONFLICT 433 433 A -> S (in Ref. 2).
SQ	SEQUENCE 500 AA; 50327 MW; E536EFFFAS212319 CRC64;
Query Match 86.6%; Score 97; DB 1; Length 500;	
Best Local Similarity 84.2%; Pred. NO. 0.18;	
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY	1 GGCGSGGGSGGGGSGGG 19 :
DB	30 GGCGGGGGGGGGAGGGG 48
RESULT 38	
ID	Q653C5 PRELIMINARY; PRT; 673 AA.
AC	Q653C5;
DT	25-OCT-2004 (TrEMBLrel. 28, Created)
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Hypothetical protein P0635GI0.21.
GN	Name=P0635GI0.21;
OS	Oryza sativa (japonica cultivar-group).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Ehrhartoideae; Oryzaceae; Oryza.
OX	NCBI_TaxID=39947;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Sasaki T., Matsumoto T., Katayose Y.;
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC clone:P0635GI0.";
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
KW	EMBL; AF005396; BAD46092.1; -.
KW	Hypothetical protein.
SQ	SEQUENCE 673 AA; 73813 MW; F99F683668802B75 CRC64;
Query Match 86.6%; Score 97; DB 2; Length 673;	
Best Local Similarity 80.0%; Pred. No. 0.23;	
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 GGCGSGGGSGGGSGGGGS 20 :
DB	77 GGCGGGGGGGGGGGGGGA 96
RESULT 39	
ID	Q91TRI PRELIMINARY; PRT; 718 AA.
ID	Q91TRI
AC	Q91TRI;

```
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE T32.
OS Tupaiid herpesvirus 1 (strain 1) (TuHV-1) (Herpesvirus tupaia (strain
OS 1)).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RX DOI=10.1128/JVI.75.10.4854-4870.2001;
RA Bahr U.; Darai G.;
RT "Analysis and characterization of the complete genome of tupaia (tree
RL shrew) herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
DR EMBL; AF281817; AAK57076.1; -.
DR Pfam; PF06070; Herpes UL32; 1.
DR PROSITE; PS00583; FPKB_KINASES_1; UNKNOWN_1.
SQ SEQUENCE 718 AA; 74589 MW; 92ED97DAC14EA58F CRC64;

Query Match 86.6%; Score 97; DB 2; Length 718;
Best Local Similarity 84.2%; Pred. No. 0.25;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 19
Db |||||:|||||
680 GGGGAGGGGGGGGGGGGGGG 698

RESULT 40
Q8PJA2
ID Q8PJA2 PRELIMINARY; PRT; 61 AA.
AC Q8PJA2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XAC2637.
GN OrderedLocusNames=XAC2637;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Cimarotte G., Cannavan F., Cardozo J., Chamberg F., Chapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011904; AAM37484.1; -.
KW Complete proteome.
SQ SEQUENCE 61 AA; 5119 MW; CCAEB43D21AFB3AD CRC64;

Query Match 85.7%; Score 96; DB 2; Length 61;
Best Local Similarity 84.2%; Pred. No. 0.035;
```

```
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGG 19
Db |||||:|||||
35 GGGGGGGGGGGGGGGGGGGG 53

Search completed: August 19, 2005, 16:35:28
Job time : 176 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:12:38 ; Search time 164 Seconds
(without alignments)

(without alignments)
11.791 Million cell updates/sec

Title: TUNGAP-CLAIM8

Perfect score: 29

Sequence: 1 gggg 5

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

```
Database :
      A_Geneseq_16Dec04.*
      1: geneseqp1980s.*
      2: geneseqp1990s.*
      3: geneseqp2000s.*
      4: geneseqp2001s.*
      5: geneseqp2002s.*
      6: geneseqp2003as.*
      7: geneseqp2003bs.*
      8: geneseqp2004s.*
```

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	29	100.0	5	6	ABG75780	hmn-14 sc	
2	29	100.0	8	5	ABB83321	Adenylate	
3	29	100.0	8	6	AAO16516	Synthetic	
4	29	100.0	10	4	ABBS6486	Human sin	
5	29	100.0	10	4	AAG94235	Human com	
6	29	100.0	10	5	ABB99179	C. elegan	
7	29	100.0	12	7	ADE29207	Multivale	
8	29	100.0	14	4	ABBS6484	Human sin	
9	29	100.0	22	2	AAR07935	Tryptic f	
10	29	100.0	32	1	AAP50011	Sequence	
11	29	100.0	35	4	AAB77627	Core poly	
12	29	100.0	35	4	ABB02109	Viral cor	
13	29	100.0	35	4	ABB00632	Viral DP1	
14	29	100.0	35	4	AAU13180	DP178-lik	
15	29	100.0	35	5	ADE02129	Hybrid po	
16	29	100.0	35	6	AAE36925	Araneus d	
17	29	100.0	38	2	AAR58503	Binding d	
18	29	100.0	39	3	AAAY4777	N-termina	
19	29	100.0	39	8	ABOS4679	Human gen	
20	29	100.0	40	5	ADK35993	Novel hum	
21	29	100.0	41	5	ADK35456	Novel hum	
22	29	100.0	41	5	ADK35544	Novel hum	
23	29	100.0	50	4	AAU64806	Propionib	
24	29	100.0	50	6	ABMG1325	Propionib	
25	29	100.0	51	5	ABP34765	Human ORF	

99 29 100.0 114 5 ABJ15132 Abj15132 Synthetic
100 29 100.0 114 8 ADM99325 Adm99325 C_heteros

ALIGNMENTS

RESULT 1
ABG75780
ID ABG75780 standard; peptide; 5 AA.

XX AC ABG75780;
XX 29-APR-2003 (first entry)
XX hMN-14 scFv VK to VH linker peptide #2.
XX Linker peptide; lesion detection; operative; intravascular; laparoscopy;
KW endoscopy; single chain antibody; scFv; tumour; lesion; catheter;
KW photoactive agent; Auger electron; ionising radiation; biopsy; cancer;
KW infectious lesion; inflammatory lesion; inflammation; clot; hyperplasia;
KW atherosclerotic plaque; brachytherapy; external beam therapy; VK; VH;
KW variable light chain; variable heavy chain; MN-14; cytostatic;
KW antiinflammatory; vulnery.

XX Synthetic.
XX US2002146369-A1.
XX 10-OCT-2002.
XX 04-MAR-2002; 2002US-00086637.
XX 06-MAY-1992; 92US-00879857.
XX 22-AUG-1994; 94US-00293313.
XX 10-FEB-1998; 98US-00021392.
XX 08-JUL-1999; 99US-00348818.
XX (IMMU-) IMMUNOMEDICS INC.
XX Goldenberg MD;
XX WFI; 2003-247116/24.

Close-range lesion detection during laparoscopy or endoscopy, by
injecting labeled divalent single chain antibody fragment that binds
antigen of lesion, scanning accessed interior of patient at close range
with detector.

Disclosure; Page 9; 18pp; English.

The invention discloses a method for close-range lesion detection during
operative, intravascular, laparoscopy or endoscopic procedure. The method
involves injecting a patient parenterally with a labelled divalent single
chain antibody (scFv) fragment, or subfragment, which binds to an antigen
produced by, or associated with, a tumour or lesion, conducting the
procedure within 48 hours of injection, scanning the accessed interior of
the patient at close range with a detector for the labelled scFv fragment
and locating the sites of accretion of the labelled scFv. Also disclosed
are methods for the treatment of lesions during a laparoscopic or
intravascular catheter procedure which involves injecting a patient with
the scFv labelled with a photoactive agent or an agent capable of being
activated to emit Auger electrons or other ionising radiation, where the
scFv accretes at the lesion, and then conducting the procedure within 48
hours of the injection, including activation of the photoactive or
ionising radiation agent, and for obtaining biopsy samples from a lesion.
The method is useful for close-range detection of lesion such as cancer,
infectious lesion, inflammatory lesion, non-tumorous or non-infectious
inflammation, clot, hyperplasia or atherosclerotic plaque, during an
operative, intravascular, laparoscopic or endoscopic procedure. The
methods do not require processing of images, both target-specific and non
-target-specific. The method enables a surgeon, through the use of an
intraoperative, laparoscopic, intravascular probe or an endoscope, to

CC scan areas of suspected tumour growth relatively quickly and use the
CC level of radiation to more precisely discriminate tumour tissue from non-
CC tumour tissue, and thereby more precisely define tumour borders for
CC surgical resection, diagnostic evaluation or for laser or radiation
CC therapy, including brachytherapy and external beam therapy or for
CC improved biopsy procedures. The sequence presented is a linker peptide
CC connecting the variable light chain (VK) to the variable heavy chain (VH)
CC of the human MN-14 scFv
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFPGS 5
| | | | |
Db 1 GGFPGS 5

RESULT 2
ABB83321
ID ABB83321 standard; peptide; 8 AA.

XX AC ABB83321;
XX 30-AUG-2002 (first entry)
XX Adenylate kinase peptide fragment #2.
DE Adenylate kinase; three dimensional structure;
KW protein structure analysis.

XX Unidentified.
XX OS
XX US2002035434-A1.
XX 21-MAR-2002.

XX 23-JUL-2001; 2001US-00910054.
XX 06-FEB-1992; 92JP-00021012.
XX 11-DEC-1992; 92JP-00331703.
XX 08-FEB-1993; 93US-00014867.

XX (FUJI) FUJITSU LTD.
XX Tomikawa M, Aikawa S, Matsuzawa F;
XX WFI; 2002-507172/54.

Analysis of three-dimensional structures by generating combination of
PT correspondence satisfying restriction condition, and calculating root
PT mean square distance between elements in the combination of
PT correspondence.

Disclosure; Fig 30; 65pp; English.

The present invention relates to a method for analysis of three
dimensional structures. The method involves generating a combination of
correspondence satisfying a restriction condition between the elements
belonging to a first and second point sets from among all candidates for
the combination of correspondence, and calculating a root mean square
distance between the elements corresponding in the combination of
correspondence. The method is useful for analysing three-dimensional
structures of molecules, particularly proteins. To illustrate the
invention, a peptide fragment of an elongation factor of a protein which
is a binding site for phosphoric acid of GTP (guanosine triphosphate),
was used as a probe sequence. The present peptide, a fragment of
adenylate kinase, was retrieved using the elongation factor peptide
Sequence 8 AA;

```

Query Match      100.0%; Score 29; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      1 GGPGS 5

RESULT 3
AAO16516
ID AAO16516 standard; peptide; 8 AA.
XX
AC AAO16516;
XX
DT 01-MAY-2003 (first entry)
XX
DE Synthetic spider silk protein linker unit repeat #5.
XX
KW Spider silk; spider silk protein; fabric; suture; medical covering;
KW high-tech clothing; rope; reinforced plastic; linker peptide.
XX
OS Unidentified.
XX
PN WO200299082-A2.
XX
PD 12-DEC-2002.
XX
PF 06-JUN-2002; 2002WO-US018256.
XX
PR 06-JUN-2001; 2001US-0296184P.
XX
PA (UWVY-) UNIV WYOMING.
XX
PI Roth DA, Lewis RV;
XX
DR WPI; 2003-140616/13.
XX
XX Expressing spider silk protein in a higher plant, by contacting a plant
PT cell with silk protein encoding a gene linked to a gene that confers
PT resistance to selection agent, and selecting cells that survive when
PT incubated with the agent.
XX
PS Example 4; Page 71; 114pp; English.
XX
XX The invention comprises a method for expressing spider silk in a higher
CC plant (e.g. arabidopsis, tobacco, tubers, sunflower, canola, alfalfa,
CC soybean, maize, sorghum, wheat, cotton, small grains and rice). The
CC method is useful for expressing spider silk in a higher plant. The silk
CC produced is useful in the production of fabrics, sutures, medical
CC coverings, high-tech clothing, rope, reinforced plastics, and other
CC applications in which various combinations of strength and elasticity are
CC required. The present amino acid sequence represents a synthetic spider
CC silk protein linker unit repeat peptide
XX
SQ Sequence 8 AA;
Query Match      100.0%; Score 29; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      3 GGPGS 7

RESULT 4
ABB56486
ID ABB56486 standard; peptide; 10 AA.
XX
AC ABB56486;
XX
XX 25-FEB-2002 (first entry)
XX

```

```

XX Human single chain MHC class II molecule linker #10.
DE
XX Human; MHC; major histocompatibility complex; MHC class II; multimer;
KW single chain; immunosuppressive; antidiabetic; antiinflammatory;
KW antianaemic; antithematoid; antiarthritic; neuroprotective; vaccine;
KW autoimmune disease; insulin dependent diabetes; multiple sclerosis;
KW myasthenia gravis; pernicious anaemia; autoimmune encephalomyelitis;
KW rheumatoid arthritis; systemic lupus erythematosus; linker.
XX
OS Homo sapiens.
XX
PN WO200170245-A1.
XX
PD 27-SEP-2001.
XX
PF 22-MAR-2001; 2001WO-US009616.
XX
PR 22-MAR-2000; 2000US-0191274P.
PR 15-MAY-2000; 2000US-0204249P.
PR 23-JAN-2001; 2001US-0264003P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Carter D, Zhu S, Arimilli S, Wang A;
XX
DR WPI; 2001-616371/71.
XX
XX Multimeric complex for treating autoimmune diseases, comprises first and
PT second single chain MHC class II molecules, each comprising alpha and
PT beta domain linked through amino acid linker and multimerization domain.
XX
PS Claim 24; Page 131; 147pp; English.
XX
XX The invention relates to a multimeric complex comprising a first
CC recombinant single chain major histocompatibility complex (MHC) class II
CC molecule and a second recombinant single chain MHC class II molecule,
CC each comprising an alpha domain and a beta domain linked through an
CC amino acid linker and a multimerisation domain. The first and the second
CC molecule are linked through the multimerisation domain to form a
CC multimeric complex. The complex is useful for treating autoimmune
CC diseases. It is useful for treating insulin dependent diabetes, multiple
CC sclerosis, myasthenia gravis, pernicious anaemia, and systemic lupus
CC encephalomyelitis (EAE), rheumatoid arthritis and autoimmune
CC erythematosus. The present sequence is a linker used to connect the
CC alpha and beta domains of the single chain MHC class II molecules of
CC the invention
XX
SQ Sequence 10 AA;
Query Match      100.0%; Score 29; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      6 GGPGS 10

RESULT 5
AAG94235
ID AAG94235 standard; peptide; 10 AA.
XX
AC AAG94235;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 429.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX

```


CC This invention relates to a novel multivalent multimeric antibody (a
 CC bispecific diabody - BbDb) which comprises at least two binding sites
 CC specific for the human B cell marker CD19 and human Pcsamra receptor III
 CC (CD16). The invention may allow the development of compositions with
 CC cytostatic or immunosuppressive activity. The antibody is useful for the
 CC diagnosis and treatment of B-cell malignancies such as non-Hodgkin
 CC lymphoma, B-cell mediated autoimmune diseases or the depletion of B-
 CC cells. The multivalent multimeric antibody avoids the undesired immune
 CC response such as human anti-murine antibody response. The specification
 CC provides a process for stable high yield of recombinant antibodies. The
 CC present sequence is that of a synthetic linker peptide which was used in
 CC the exemplification of the invention.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 29; DB 7; Length 12;

Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5

Db 8 GPGGS 12

RESULT 8

ABB56484

ID ABB56484 standard; peptide; 14 AA.

XX AC ABB56484;

XX DT 25-FEB-2002 (first entry)

XX DE Human single chain MHC class II molecule linker #8.

XX KW Human; MHC; major histocompatibility complex; MHC class II; multimer;
 XX KW single chain; immunosuppressive; antidiabetic; anti-inflammatory;
 XX KW antianemic; antirheumatoid; antithyroidic; neuroprotective; vaccine;
 XX KW autoimmune disease; insulin dependent diabetes; multiple sclerosis;
 XX KW myasthenia gravis; pernicious anemia; autoimmune encephalomyelitis;
 XX KW rheumatoid arthritis; systemic lupus erythematosus; linker.

XX OS Homo sapiens.

XX PN WO200170245-A1.

XX PD 27-SEP-2001.

XX PF 22-MAR-2001; 2001WO-US009616.

XX PR 22-MAR-2000; 2000US-0191274P.

XX PR 15-MAY-2000; 2000US-0204249P.

XX PR 23-JAN-2001; 2001US-0264003P.

XX PA (CORI-) CORIXA CORP.

XX PI Carter D, Zhu S, Arimilli S, Wang A;

XX DR WPI; 2001-616371/71.

XX PT Multimeric complex for treating autoimmune diseases, comprises first and
 XX PT second single chain MHC class II molecules, each comprising alpha and
 XX PT beta domain linked through amino acid linker and multimerization domain.

XX PS Claim 24; Page 131; 147pp; English.

XX CC The invention relates to a multimeric complex comprising a first
 XX CC recombinant single chain major histocompatibility complex (MHC) class II
 XX CC molecule and a second recombinant single chain MHC class II molecule,
 XX CC each comprising an alpha domain and a beta domain linked through an
 XX CC amino acid linker and a multimerization domain. The first and the second
 XX CC molecule are linked through the multimerization domain to form a
 XX CC multimeric complex. The complex is useful for treating autoimmune
 XX CC diseases. It is useful for treating insulin dependent diabetes, multiple

CC sclerosis, myasthenia gravis, pernicious anaemia, autoimmune
 CC encephalomyelitis (EAE), rheumatoid arthritis and systemic lupus
 CC erythematosus. The present sequence is a linker used to connect the
 CC alpha and beta domains of the single chain MHC class II molecules of
 CC the invention

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 29; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5

Db 10 GPGGS 14

RESULT 9

AAR07935

ID AAR07935 standard; protein; 22 AA.

XX AC AAR07935;

XX DT 25-MAR-2003 (revised)

XX DT 14-JAN-1991 (first entry)

XX DE Tryptic fragment T31 of rat phospholipase A2 inhibitor.

XX KW Human N-lipocortin; placenta; inflammation reduction; arthritis;

XX KW rat phospholipase A2 inhibitor; tryptic fragment T31.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT Misc-difference 12..12 /label= S, T

XX FT Misc-difference 15..15 /label= S, T

XX PN US4950646-A.

XX PD 21-AUG-1990.

XX PF 10-JAN-1986; 86US-00929199.

XX PR 10-JAN-1985; 85US-00690146.

XX PR 15-MAR-1985; 85US-00712376.

XX PR 14-AUG-1985; 85US-00765877.

XX PR 05-SEP-1985; 85US-00772892.

XX PA (BIOJ) BIOGEN INC.

XX PI Wallner BP, Pepinsky RB, Garwin JL, Schindler DG, Huang KS;

XX DR WPI; 1990-274549/36.

XX PT Pure fragment of human lipocortin - useful for reducing inflammation or
 XX PT for treating arthritis, etc.

XX PS Disclosure; Fig 2; 51pp; English.

XX CC Rat phospholipase A2 inhibitor protein was isolated from the
 XX CC extracellular supernatant of rat peritoneal exudate cells. The protein
 XX CC was isolated and sequenced to produce oligonucleotide probes in order to
 XX CC identify human lipocortin and N-lipocortin. See also AAQ05805-25,
 XX CC AAQ06581, AAR0926-37 and AAR07956-66. (Updated on 25-MAR-2003 to correct
 XX CC PA field.)

XX SQ Sequence 22 AA;

Query Match 100.0%; Score 29; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GGFGS 5
Db      1 GGFGS 5

RESULT 10
AAP50011 standard; peptide; 32 AA.
XX
AC AAP50011;
XX
AC      (revised)
DT 25-MAR-2003
DT 16-AUG-2002 (revised)
DT 03-SEP-1991 (first entry)
XX
Sequence linking insulin B chain B(1-29) to A chain a(1-21) to form a
DE biosynthetic insulin precursor.
DE
XX Insulin precursor; diabetes; proteolysis resistant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PH      Location/Qualifiers
FT Misc-difference 1 /note= "attached to B(1-29)"
FT FT
FT Misc-difference 34
FT FT
FT /note= "attached to A(1-21)"
XX
XX
PN EP163529-A.
XX
XX 04-DEC-1985.
XX
XX 30-MAY-1984; 84DK-00002665.
XX
XX 30-MAY-1984; 84DK-00002665.
PR 08-FEB-1985; 85DK-00000582.
PR 29-MAY-1985; 85DK-00002385.
XX
XX (NOVO ) NOVO IND AS.
XX
XX Markussen J, Fiil N, Ammerer G, Hansen M, Thim L, Norris K;
PI Voigt HO;
XX
XX WPI; 1985-304970/49.
XX
XX DNA-sequence encoding insulin precursor - having correctly positioned
PT disulphide bridges and resistant to proteolytic degradation.
XX
XX Disclosure; Page 11; 47pp; English.
XX
XX The biosynthetic insulin precursors are generated largely with correctly
CC positioned disulphide bridges between the A- and B- moieties, and are
CC more resistant to proteolytic degradation than previously. (Updated on 16
CC -AUG-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
XX
SQ      Sequence 32 AA;
Query Match      100.0%; Score 29; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGS 5
Db      16 GGFGS 20

RESULT 11
AAB77627
ID AAB77627 standard; peptide; 35 AA.
XX

```

```

AC AAB77627;
XX
DT 19-APR-2001 (first entry)
XX
DE Core polypeptide T713.
XX
KW Core polypeptide; enhancer; antiviral; anti-HIV; virucide; hepatotropic;
KW antiinflammatory; hybrid polypeptide; coiled-coil peptide interaction;
KW fusion-related disorder; bacterial infection; viral infection.
XX
OS Unidentified.
XX
PN WO200103723-A1.
XX
PD 18-JAN-2001.
XX
PF 10-JUL-2000; 2000WO-US018772.
XX
PR 09-JUL-1999; 99US-00350641.
XX
PA (TRIM-) TRIMERIS INC.
XX
PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
XX
XX WPI; 2001-147136/15.
XX
XX New hybrid polypeptide, useful for preventing, treating and diagnosing
PT e.g. viral infections, comprises an enhancer peptide linked to a core
PT polypeptide.
XX
PS Disclosure; Page 44; 151pp; English.
XX
XX The present sequence is a core polypeptide which may be linked to an
CC enhancer peptide to form a novel hybrid polypeptide. The hybrid
CC polypeptide exhibits enhanced pharmacokinetic properties relative to
CC those exhibited by the core polypeptide when introduced into a living
CC system. It is used to increase the in vitro or ex vivo half-life of the
CC core polypeptide. The hybrid and core polypeptides can be used for
CC modulating fusogenic events and intracellular processes involving coiled-
CC coil peptide interactions. Other uses include preventing, treating and/or
CC diagnosing disorders involving fusion events (e.g. modulation of
CC neurotransmitter exchange and sperm-egg fusion), intracellular processes
CC involving coiled-coil peptides (e.g. bacterial infections) and viral
CC infections that involve cell-cell and/or virus-cell fusion (e.g. viral
CC infections caused by human immunodeficiency virus, respiratory syncytial
CC virus, Epstein-Barr virus, hepatitis B virus, Mason-Pfizer virus and
CC polio virus). The enhancer peptide sequence increases the half-life and
CC reduces the clearance rate of therapeutic peptides, which increases their
CC efficacy and minimises the incidence and severity of adverse side
CC effects. In addition, this increases the sensitivity of the diagnostic
CC procedure in which they are used
XX
SQ      Sequence 35 AA;
Query Match      100.0%; Score 29; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGS 5
Db      14 GGFGS 18

RESULT 12
ABB02109
ID ABB02109 standard; peptide; 35 AA.
XX
XX ABB02109;
AC
XX
XX 11-SEP-2003 (revised)
DT 06-AUG-2003 (revised)
DT 03-JAN-2002 (first entry)
XX

```


PF 05-JUL-2000; 2000WO-US035727.
 XX
 PR 09-JUL-1999; 99US-00350841.
 XX
 PA (TRIM-) TRIMERIS INC.
 XX
 PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
 XX
 DR WPI; 2001-442157/47.
 XX
 XX Identifying a compound that inhibits the formation of or disrupts a
 PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
 PT or intracellular modulatory activity, by detecting the formation of a
 PT DP107/DP178 complex.
 XX
 PS Disclosure; Page 65; 259pp; English.
 XX
 CC The present invention relates to peptides which exhibit anti-retroviral
 CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
 CC DP178-like and DP107-like peptides. The DP178 peptide corresponds to
 CC amino acids 639-673 of the transmembrane protein gp41 from human
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
 CC also relates to a method of identifying compounds that inhibit the
 CC formation of or disrupts a DP107/DP178 complex. The method comprises
 CC detecting the formation of a DP107/DP178 complex, both in the presence or
 CC absence of a test compound, in a reaction mixture containing DP107 and
 CC DP178 peptides. The method is useful for identifying compounds, including
 CC small molecule compounds, which may themselves exhibit antifusogenic,
 CC antiviral or intracellular modulatory activity. The DP178-like/DP107-like
 CC peptides are useful to inhibit human and non-human retroviral,
 CC particularly HIV, transmission to uninfected cells. The present sequence
 CC represents one of the DP178-like/DP107-like peptides of the invention
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 29; DB 4; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GPGGS 5
 DB 14 GPGGS 18
 RESULT 15
 ADE02129
 ID ADE02129 standard; peptide; 35 AA.
 XX
 AC ADE02129;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Hybrid polypeptide pharmacokinetic enhancer peptide, SEQ ID No 636.
 XX
 KW hybrid; enhancer; anti-fusogenic; antiviral; virucide; antidiabetic;
 KW pharmacokinetic; fusogenic; insulin; diabetes.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Residue is modified by acetyl group"
 FT Modified-site 35
 FT /note= "C-terminal amide"
 XX
 EN US6348568-B1.
 XX
 PD 19-FEB-2002.
 XX
 PF 20-MAY-1999; 99US-00315304.
 XX
 XX 20-MAY-1998; 98US-00082279.
 PR

XX (TRIM-) TRIMERIS INC.
 PA Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
 XX
 PI WPI; 2002-424396/45.
 XX
 DR New hybrid polypeptide for modulating fusogenic events for e.g. antiviral
 XX activity, has enhancer peptide sequence derived from retroviral envelope
 PT protein sequences linked to core polypeptide e.g. therapeutic protein.
 PT
 XX Disclosure; SEQ ID NO 636; 70pp; English.
 PS
 CC The invention relates to a novel hybrid polypeptide comprising an
 CC enhancer peptide sequence linked to a core polypeptide. The enhancer
 CC peptide sequence comprises WOEWQKI or WASLWQWF. The invention also
 CC includes novel peptides that exhibit anti-fusogenic activity, antiviral
 CC activity and/or ability to modulate intracellular processes. The novel
 CC hybrid polypeptide has virucide and antidiabetic activity. The enhancer
 CC peptide sequence enhances pharmacokinetic properties of any core
 CC polypeptide, for example, a polypeptide useful for the treatment or
 CC prevention of a disease, or an imaging agent useful for imaging
 CC structures in vivo. The core polypeptides and hybrid polypeptides are
 CC useful for modulating fusogenic events and exhibit antifusogenic or
 CC antiviral activity. The novel hybrid polypeptide is useful for decreasing
 CC viral infection and modulating intracellular processes involving coiled-
 CC coil peptide interactions. The novel hybrid polypeptide comprises insulin
 CC or its fragment, so the core polypeptide is useful for ameliorating the
 CC symptoms of forms of diabetes. The novel hybrid polypeptide is also
 CC useful as a part of prognosis for preventing disorders including fusion
 CC events and viral infection that involves cell-cell and/or virus-cell
 CC fusion, and for diagnosis and in vivo imaging methods. This sequence
 CC represents an enhancer peptide of the invention.
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 29; DB 5; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GPGGS 5
 DB 14 GPGGS 18
 RESULT 16
 AAE36925
 ID AAE36925 standard; protein; 35 AA.
 XX
 AC AAE36925;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DE Araneus diadematus major ampullate spidroin 2 consensus protein #2.
 XX
 KW Spider silk protein; flagelliform-like spider silk protein; FLAG; MaSp;
 KW major ampullate spidroin; fabric; medical covering; clothing; parachute;
 KW rope; reinforced plastic; suture; aerial device wing; sail; hand glider;
 KW implantation biomaterial; resin product; fiber-reinforced plastic;
 KW thermal injected plastic.
 XX
 OS Argiope diadematus.
 XX
 PN WO2003020916-A2.
 XX
 PD 13-MAR-2003.
 XX
 XX 28-MAR-2002; 2002WO-US009663.
 PF
 XX 29-AUG-2001; 2001US-0315529P.
 PR
 PA (UYWY-) UNIV WYOMING.
 XX


```

DT XX 06-MAY-2004 (first entry)
DE XX Novel human polypeptide SeqID8075.
DX XX
KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;
KW immunosuppressive; cytostatic; antiparasitic; antiinflammatory;
KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
KW fungus; parasite; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..40
FT /label= OTHER
FT /note= "OTHER= All Xaa's in this sequence are unknown
FT amino acids or the site of a stop codon within the DNA
FT sequence"
XX
XX WO200216439-A2.
XX
XX 28-FEB-2002.
XX
XX 05-MAR-2001; 2001WO-US004941.
XX
XX 07-MAR-2000; 2000US-00519705.
XX
XX 19-MAY-2000; 2000US-00574454.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2002-280918/32.
XX
XX Isolated polynucleotide encoding bone marrow derived polypeptides useful
XX for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
XX disease, and inflammatory bowel disease.
XX
XX Claim 20; SEQ ID NO 8075; 504pp; English.
XX
XX This invention relates to a novel isolated polynucleotide comprising a
XX nucleotide sequence selected from one of 1680 sequences, a mature protein
XX coding portion of them, an active domain of them and their complementary
XX sequences. The invention may be useful for the production of compounds
XX with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
XX immunosuppressive, cytostatic, antiparasitic, antiinflammatory,
XX antibacterial, antiviral, antifungal or antiparasitic activity. In
XX addition, the disclosed sequences may be useful for gene therapy. The
XX polypeptides or their antibodies are useful for treating many diseases
XX such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
XX psoriasis, inflammatory bowel disease and infections caused by bacteria,
XX viruses, fungi or parasites. The present sequence is that of a human
XX polypeptide of the invention.
XX
XX Sequence 40 AA;
SQ
Query Match 100.0%; Score 29; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GPGGS 5
DB 19 GPGGS 23
RESULT 21
ADK35456
ID ADK35456 standard; protein; 41 AA.
XX
XX AC ADK35456;
XX
XX DT 06-MAY-2004 (first entry)

```

```

XX Novel human polypeptide SeqID7538.
XX
XX antiarthritic; antiparkinsonian; neuroprotective; nootropic;
XX immunosuppressive; cytostatic; antiparasitic; antiinflammatory;
XX antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
XX arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
XX psoriasis; inflammatory bowel disease; infection; bacteria; virus;
XX fungus; parasite; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..41
FT /label= OTHER
FT /note= "OTHER= All Xaa's in this sequence are unknown
FT amino acids or the site of a stop codon within the DNA
FT sequence"
XX
XX WO200216439-A2.
XX
XX 28-FEB-2002.
XX
XX 05-MAR-2001; 2001WO-US004941.
XX
XX 07-MAR-2000; 2000US-00519705.
XX
XX 19-MAY-2000; 2000US-00574454.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2002-280918/32.
XX
XX Isolated polynucleotide encoding bone marrow derived polypeptides useful
XX for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
XX disease, and inflammatory bowel disease.
XX
XX Claim 20; SEQ ID NO 7538; 504pp; English.
XX
XX This invention relates to a novel isolated polynucleotide comprising a
XX nucleotide sequence selected from one of 1680 sequences, a mature protein
XX coding portion of them, an active domain of them and their complementary
XX sequences. The invention may be useful for the production of compounds
XX with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
XX immunosuppressive, cytostatic, antiparasitic, antiinflammatory,
XX antibacterial, antiviral, antifungal or antiparasitic activity. In
XX addition, the disclosed sequences may be useful for gene therapy. The
XX polypeptides or their antibodies are useful for treating many diseases
XX such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
XX psoriasis, inflammatory bowel disease and infections caused by bacteria,
XX viruses, fungi or parasites. The present sequence is that of a human
XX polypeptide of the invention.
XX
XX Sequence 41 AA;
SQ
Query Match 100.0%; Score 29; DB 5; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GPGGS 5
DB 18 GPGGS 22
RESULT 22
ADK35544
ID ADK35544 standard; protein; 41 AA.
XX
XX AC ADK35544;
XX
XX DT 06-MAY-2004 (first entry)

```

DE Novel human polypeptide SeqID7626.

XX antiarthritic; antiparkinsonian; neuroprotective; nontropic;
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
 KW fungus; parasite; human.

XX Homo sapiens.

XX Key Location/Qualifiers
 PH Misc-difference 1: .41
 FT /label= OTHER
 FT /note= "OTHER= All Xaa's in this sequence are unknown
 FT amino acids or the site of a stop codon within the DNA
 FT sequence"

XX WO200216439-A2.

XX 28-FEB-2002.

XX 05-MAR-2001; 2001WO-US004941.

XX 07-MAR-2000; 2000US-00519705.

XX 19-MAY-2000; 2000US-00574454.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2002-280918/32.

XX Isolated polynucleotide encoding bone marrow derived polypeptides useful
 PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 PT disease, and inflammatory bowel disease.

XX Claim 20; SEQ ID NO 7626; 504pp; English.

XX This invention relates to a novel isolated polynucleotide comprising a
 CC nucleotide sequence selected from one of 1680 sequences, a mature protein
 CC coding portion of them, an active domain of them and their complementary
 CC sequences. The invention may be useful for the production of compounds
 CC with an antiarthritic, antiparkinsonian, neuroprotective, nontropic,
 CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory, In
 CC antibacterial, antiviral, antifungal or antiparasitic activity. In
 CC addition, the disclosed sequences may be useful for gene therapy. The
 CC polypeptides or their antibodies are useful for treating many diseases
 CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
 CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
 CC viruses, fungi or parasites. The present sequence is that of a human
 CC polypeptide of the invention.

XX Sequence 41 AA;

Query Match 100.0%; Score 29; DB 5; Length 41;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
 DB 18 GGPGS 22

RESULT 23
 AAU64806
 ID AAU64806 standard; protein; 50 AA.
 XX AAU64806;
 AC AAU64806;
 XX 27-FEB-2002 (first entry)
 DT
 XX Propionibacterium acnes immunogenic protein #25702.

XX Novel human polypeptide SeqID7626.

XX antiarthritic; antiparkinsonian; neuroprotective; nontropic;
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
 KW fungus; parasite; human.

XX Homo sapiens.

XX Key Location/Qualifiers
 PH Misc-difference 1: .41
 FT /label= OTHER
 FT /note= "OTHER= All Xaa's in this sequence are unknown
 FT amino acids or the site of a stop codon within the DNA
 FT sequence"

XX WO200216439-A2.

XX 28-FEB-2002.

XX 05-MAR-2001; 2001WO-US004941.

XX 07-MAR-2000; 2000US-00519705.

XX 19-MAY-2000; 2000US-00574454.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2002-280918/32.

XX Isolated polynucleotide encoding bone marrow derived polypeptides useful
 PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 PT disease, and inflammatory bowel disease.

XX Claim 20; SEQ ID NO 7626; 504pp; English.

XX This invention relates to a novel isolated polynucleotide comprising a
 CC nucleotide sequence selected from one of 1680 sequences, a mature protein
 CC coding portion of them, an active domain of them and their complementary
 CC sequences. The invention may be useful for the production of compounds
 CC with an antiarthritic, antiparkinsonian, neuroprotective, nontropic,
 CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory, In
 CC antibacterial, antiviral, antifungal or antiparasitic activity. In
 CC addition, the disclosed sequences may be useful for gene therapy. The
 CC polypeptides or their antibodies are useful for treating many diseases
 CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
 CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
 CC viruses, fungi or parasites. The present sequence is that of a human
 CC polypeptide of the invention.

XX Sequence 41 AA;

Query Match 100.0%; Score 29; DB 5; Length 41;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
 DB 18 GGPGS 22

RESULT 23
 AAU64806
 ID AAU64806 standard; protein; 50 AA.
 XX AAU64806;
 AC AAU64806;
 XX 27-FEB-2002 (first entry)
 DT
 XX Propionibacterium acnes immunogenic protein #25702.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.

XX N-PSDB; AAS59652.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX Example 1; SEQ ID NO 26001; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 50 AA;

Query Match 100.0%; Score 29; DB 4; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
 DB 20 GGPGS 24

RESULT 24
 ABM61325
 ID ABM61325 standard; protein; 50 AA.
 XX ABM61325;
 AC ABM61325;
 XX 20-OCT-2003 (first entry)
 DT
 XX Propionibacterium acnes predicted ORF-encoded polypeptide #26001.

XX Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX Propionibacterium acnes.
 XX WO2003033515-A1.
 XX 24-APR-2003.
 XX 11-OCT-2002; 2002WO-US032727.
 XX 15-OCT-2001; 2001US-00978825.
 XX (CORI-) CORIYA CORP.
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglas J;
 XX WPI; 2003-381789/36.
 DR N-PSDB; ACF64581.
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX Example 1; SEQ ID NO 26001; 1481pp; English.
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 50 AA;
 Query Match 100.0%; Score 29; DB 6; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGPGS 5
 Db 20 GGGPGS 24
 |||||
 RESULT 25
 ABP34765
 ID ABP34765 standard; protein; 51 AA.
 XX AC ABP34765;

XX 08-JUL-2002 (first entry)
 DE Human ORF3738 protein, SEQ ID NO:7476.
 XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX Homo sapiens.
 XX WO200190366-A2.
 XX 29-NOV-2001.
 XX 24-MAY-2001; 2001WO-US017076.
 XX 24-MAY-2000; 2000US-0206690P.
 XX (CURA-) CURAGEN CORP.
 XX Leach MD, Shinkets RA;
 XX WPI; 2002-106200/14.
 XX N-PSDB; ABN78791.
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation.
 XX Claim 10; Page 2115; 2508pp; English.
 XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC of both infective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals

CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases
 XX
 SQ Sequence 51 AA;

Query Match 100.0%; Score 29; DB 5; Length 51;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
 DB 25 GGPGS 29
 |||||

RESULT 26
 ADJ67007
 ID ADJ67007 standard; protein; 53 AA.

AC ADJ67007;
 XX
 XX 06-MAY-2004 (first entry)
 XX Human secreted protein (SECP) #33.
 XX human; secreted protein; SECP; cardiovascular disorder;
 KW neurological disorder; cancer.
 XX Homo sapiens.
 XX WO2004007527-A2.

XX 22-JAN-2004.
 XX 10-JUL-2003; 2003WO-US021576.
 XX 11-JUL-2002; 2002US-0395722P.
 PR 26-JUL-2002; 2002US-0395003P.
 PR 01-AUG-2002; 2002US-0400909P.
 PR 30-AUG-2002; 2002US-0407202P.
 PR 30-AUG-2002; 2002US-0407552P.

XX (INCY-) INCYTE CORP.
 XX Kable AE, Swarnakar A, Lee SY, Richardson TW, Ison CH, Khare R;
 PI Jin P, Fu GK, Becha SD, Tang YT, Elliott VS, Warren BA, Jiang X;
 PI Jackson AA, Baughn MR, Duggan BM, Hafalia AJA, Ramkumar J, Tran UK;
 PI Lindquist EA, Chawla NK, Marquis JP, Wang JT;
 XX WPI; 2004-122875/12.
 DR N-PSDB; ADJ67061.

XX New human secreted proteins (SECP), useful for preparing a composition
 PT for diagnosing or treating a disease or condition associated with
 PT decreased expression or over expression of functional SECP e.g. cancer.

XX Claim 1; SEQ ID NO 33; 285pp; English.

XX The invention comprises the amino acid and coding sequences of human
 CC secreted proteins (SECP). The DNA and protein sequences of the invention
 CC are useful for diagnosing or treating a disease or condition associated
 CC with decreased expression or over expression of functional SECP, such as:
 CC cardiovascular disorders, neurological disorders and cancer. The present
 CC amino acid sequence represents a human SECP of the invention.

XX Sequence 53 AA;

Query Match 100.0%; Score 29; DB 8; Length 53;
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5

DB 28 GGPGS 32
 |||||

RESULT 27
 AAU62800
 ID AAU62800 standard; protein; 54 AA.

XX AAU62800;
 AC AAU62800;
 XX 27-FEB-2002 (first entry)
 XX Propionibacterium acnes immunogenic protein #23696.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59629.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX Example 1; SEQ ID NO 23995; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 54 AA;

Query Match 100.0%; Score 29; DB 4; Length 54;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5

Db 12 GPGS 16

|||||

RESULT 28

ABM59319

ID ABM59319 standard; protein; 54 AA.

XX

AC ABM59319;

XX

DT 20-OCT-2003 (first entry)

XX

DE Propionibacterium acnes predicted ORF-encoded polypeptide #23995.

XX

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.

KW

XX

OS Propionibacterium acnes.

XX

XX

PN WO2003033515-A1.

XX

XX

PD 24-APR-2003.

XX

XX 11-OCT-2002; 2002WO-US032727.

PF

XX

XX 15-OCT-2001; 2001US-00978825.

PR

XX

XX (CORI-) CORIXA CORP.

PA

XX

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieue-Douglass J;

XX

XX

DR WPI; 2003-381789/36.

DR N-PSDB; ACF64558.

XX

XX

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX

XX

PS Example 1; SEQ ID NO 23995; 1481pp; English.

XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC encoding a Propionibacterium acnes protein. The invention also relates to

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of P. acnes polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the

CC invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes

CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or

CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of P. acnes in a

CC patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne

CC vulgaris, or for stimulating an immune response specific for a P. acnes

CC protein. The polynucleotides can also be used as probes or primers for

CC nucleic acid hybridisation. The vaccine composition is useful for the

CC stimulation of an immune response against P. acnes, or for treating acne,

CC and the kit is useful for performing a diagnostic assay. The present

CC sequence represents a polypeptide predicted to be encoded by an ORF (open

CC reading frame) contained within the P. acnes polynucleotides of the

CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/obtained_pct_sequences

XX

SQ Sequence 54 AA;

Query Match 100.0%; Score 29; DB 6; Length 54;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5

Db 12 GPGS 16

|||||

RESULT 29

ABP03331

ID ABP03331 standard; protein; 55 AA.

XX

XX ABP03331;

XX

DT 25-JUN-2002 (first entry)

XX

DE Human ORFX protein sequence SEQ ID NO:6644.

XX

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;

KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;

KW degenerative disorder; osteoarthritis; neurodegenerative disorder;

KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;

KW hypertension; hypothyroidism; cholesterol ester storage disease;

KW immune deficiency; immune disorder; infectious disease;

KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;

XX myasthenia gravis.

XX

OS Homo sapiens.

XX

XX WO200192523-A2.

PN

XX

PD 06-DEC-2001.

PD

XX

XX 29-MAY-2001; 2001WO-US010836.

PF

XX

XX 30-MAY-2000; 2000US-0206132P.

PR

XX

XX 29-AUG-2000; 2000US-0228716P.

PR

XX

XX (CURA-) CURAGEN CORP.

PA

XX

PI Shimkets RA, Leach MD;

PI

XX

DR WPI; 2002-106308/14.

DR N-PSDB; ABN19083.

DR

XX

PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,

PT hyperproliferative disorders and autoimmune disorders.

XX

PS Disclosure; SEQ ID NO 6644; 1037pp; English.

XX

CC The present invention describes substantially purified human proteins

CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX

CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for

CC treating or preventing a pathology associated with an ORFX-associated

CC disorder in humans, and in the manufacture of a medicament for treating a

CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide

CC sequences can be used in gene therapy. ORFX sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,

CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,

CC osteoarthritis, neurodegenerative disorders, disorders related to organ

CC transplantation, cardiovascular diseases, diabetes mellitus, systemic

CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

CC storage disease, various immune deficiencies and disorders, infectious

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid

CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host

CC disease and autoimmune inflammatory eye disease. ORFX proteins are also

CC useful for treating burns, incisions, ulcers, for treating osteoporosis,

CC bone degenerative disorders, or periodontal disease, and for gut

CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues and conditions resulting from

CC systemic cytokine damage. N.B. The sequence data for this patent did not

```
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 55 AA;

Query Match          100.0%; Score 29; DB 5; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
DB |||||
9 GPGGS 13

RESULT 30
ABB40703
ID ABB40703 standard; peptide; 56 AA.
XX
XX ABB40703;
AC
XX
XX 04-FEB-2002 (first entry)
DT
XX
XX Peptide #8209 encoded by human foetal liver single exon probe.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157277-A2.
FN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000669.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX
XX 26-MAY-2000; 2000US-0207456P.
PR
XX
XX 30-JUN-2000; 2000US-00608408.
PR
XX
XX 03-AUG-2000; 2000US-00632366.
PR
XX
XX 21-SEP-2000; 2000US-0234687P.
PR
XX
XX 27-SEP-2000; 2000US-0236359P.
PR
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-483447/52.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX
XX Claim 27; SEQ ID NO 33338; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 56 AA;
QY Query Match          100.0%; Score 29; DB 4; Length 56;
DB Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
DB |||||
26 GPGGS 30

Query Match          100.0%; Score 29; DB 4; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
DB |||||
26 GPGGS 30
```

```
RESULT 31
AAM34466
ID AAM34466 standard; protein; 56 AA.
XX
XX AAM34466;
AC
XX
XX 17-OCT-2001 (first entry)
DT
XX
XX Peptide #8503 encoded by probe for measuring placental gene expression.
DE
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200157272-A2.
FN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000663.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX
XX 26-MAY-2000; 2000US-0207456P.
PR
XX
XX 30-JUN-2000; 2000US-00608408.
PR
XX
XX 03-AUG-2000; 2000US-00632366.
PR
XX
XX 21-SEP-2000; 2000US-0234687P.
PR
XX
XX 27-SEP-2000; 2000US-0236359P.
PR
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488897/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 34735; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
XX Sequence 56 AA;
QY Query Match          100.0%; Score 29; DB 4; Length 56;
DB Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
DB |||||
26 GPGGS 30

RESULT 32
AAM74353
ID AAM74353 standard; protein; 56 AA.
XX
XX AAM74353;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 34659.
DE
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX
```



```

PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 34659; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention
XX
XX Sequence 56 AA;
XX
XX Query Match 100.0%; Score 29; DB 4; Length 56;
XX Best Local Similarity 100.0%; Pred. NO. 1.6e+03;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GPGS 5
XX Db 26 GPGS 30
XX
XX
XX RESULT 33
XX AAU41678
XX ID AAU41678 standard; protein; 56 AA.
XX
XX AC AAU41678;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Propionibacterium acnes immunogenic protein #2574.
XX
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX OS Propionibacterium acnes.
XX
XX PN WO200181581-A2.
XX
XX PD 01-NOV-2001.
XX
XX PF 20-APR-2001; 2001WO-US012865.
XX
XX PR 21-APR-2000; 2000US-019047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

```

```

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59515.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 2873; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 56 AA;
XX
XX Query Match 100.0%; Score 29; DB 4; Length 56;
XX Best Local Similarity 100.0%; Pred. NO. 1.6e+03;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GPGS 5
XX Db 49 GPGS 53
XX
XX
XX RESULT 34
XX ABG56151
XX ID ABG56151 standard; peptide; 56 AA.
XX
XX AC ABG56151;
XX
XX DT 25-FEB-2003 (first entry)
XX
XX DE Human liver peptide, SEQ ID No 34799.
XX
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX hypercholesterolaemia; coronary heart disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000664.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

```

```

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI gene expression in human adult liver.
XX WPI; 2001-48898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX Claim 27; SEQ ID NO 34799; 658pp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 56 AA;
SQ
Query Match 100.0%; Score 29; DB 4; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GPGGS 5
DB 26 GPGGS 30
RESULT 35
ABP08854
ID ABP08854 standard; protein; 56 AA.
XX AC ABP08854;
XX 25-JUN-2002 (first entry)
XX Human ORFX protein sequence SEQ ID NO:17690.
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX Homo sapiens.
XX WO200192523-A2.
XX 06-DEC-2001.
XX 29-MAY-2001; 2001WO-US010836.
XX 30-MAY-2000; 2000US-0206132P.
XX 29-AUG-2000; 2000US-0228716P.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach MD;
XX WPI; 2002-106308/14.
XX N-PSDB; ABN24606.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX Disclosure; SEQ ID NO 17690; 1037pp; English.
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC syndrome in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, rheumatoid
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 56 AA;
SQ
Query Match 100.0%; Score 29; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GPGGS 5
DB 39 GPGGS 43
RESULT 36
ABM38197
ID ABM38197 standard; protein; 56 AA.
XX AC ABM38197;
XX 20-OCT-2003 (first entry)
XX Propionibacterium acnes predicted ORF-encoded polypeptide #2873.
XX Propionibacterium acnes
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX Propionibacterium acnes.
XX WO2003033515-A1.
XX 24-APR-2003.
XX 11-OCT-2002; 2002WO-US032727.
XX 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barth B, Valliave-Douglass J;
XX WPI; 2003-381789/36.
XX N-PSDB; ACF64444.
XX

```

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 2873; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC polynucleotide; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 56 AA;

Query Match 100.0%; Score 29; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db 49 GPGGS 53

RESULT 37
ABP35183
ID ABP35183 standard; protein; 59 AA.

XX ABP35183;

AC 08-JUL-2002 (first entry)

DE Human ORF4156 protein, SEQ ID NO:8312.

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neutronic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.

XX WO200190366-A2.

XX

PD 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US017076.

PF 24-MAY-2000; 2000US-0206690P.

XX (CURA-) CURAGEN CORP.

XX Leach MD, Shinkets RA;

XX WPI; 2002-106200/14.

XX N-PSDB; ABN79209.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX
XX Claim 10; Page 2322; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases

XX Sequence 59 AA;

Query Match 100.0%; Score 29; DB 5; Length 59;

Best Local Similarity 100.0%; Pred. No. 1.7e+03;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5

Db 51 GPGGS 55

RESULT 38

AAW62243

ID AAW62243 standard; protein; 60 AA.

XX

AC AAW62243;

XX

DT 23-SEP-1998 (first entry)
 XX Subtilase BSPPF fragment.
 XX Subtilase; variant; mutant; subtilisin; serine protease; *Bacillus* sp;
 KW auto-proteolytic stability; laundry; dishwashing; proteolytic activity;
 KW detergent.
 XX *Bacillus subtilis*.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 21 /label= unknown
 FT /note= "X represents a gap in the sequence corresponding
 FT to residues 146-155 when in alignment with BASBPB"
 FT Misc-difference 41
 FT /label= unknown
 FT /note= "X represents a gap in the sequence corresponding
 FT to residues 173-188 when in alignment with BASBPB"
 XX
 XX WO9820116-A1.
 XX
 XX 14-MAY-1998.
 XX
 XX 04-NOV-1997; 97WO-DK000500.
 XX
 XX 04-NOV-1996; 96DK-00001235.
 PR 05-NOV-1996; 96DK-00001240.
 PR 14-MAR-1997; 97DK-00000284.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 XX
 XX Von Der Osten C, Halkier T, Andersen C, Bauditz P, Hansen PK;
 XX WPI; 1998-286932/25.
 XX
 XX New subtilase variants with increased auto-proteolytic stability and
 PT related nucleic acid, vectors and transformed cells - useful in laundry
 PT and dishwashing compositions, having increased proteolytic activity.
 XX
 XX Disclosure; Fig 1; 99pp; English.
 PS
 XX
 CC New subtilase enzyme variants have been derived from a precursor
 CC subtilase having either (i) an autoproteolytic cleavage site between
 CC residues 132 and 133 (all numbering according to the subtilisin BPN
 CC sequence (BASBPB)) or (ii) at least 1 amino acid (aa) modification at
 CC positions corresponding to G159, S164, I165, Y167, R170 or Y171 of
 CC BLS309, by substitution, insertion or deletion of at least 1 of the
 CC residues 129-136 so that autoproteolysis stability is increased, relative
 CC to unmodified enzyme. Subtilase variants are formulated in detergent
 CC compositions together with a cellulase, lipase, cutinase, oxidoreductase,
 CC another protease or amylase. The other components are e.g. standard
 CC surfactants, builders or bleaches. A typical granular fabric cleaning
 CC composition comprises 6.5% sodium 12C alkylbenzenesulphonate; 15% Na2
 CC sulphate; 26% zeolite A; 5% Na nitrilotriacetate; 0.1% of the subtilase
 CC variant; 0.5% poly(vinyl- pyrrolidone); 3% tetraacetyl ethylenediamine;
 CC 4% boric acid; 18% Na perborate monohydrate; 0.1% phenol sulphonate and
 CC minor components to make 100%. The subtilase variants are used in laundry
 CC and dishwashing detergents also in fabric softeners and hard-surface
 CC cleaners. Increased resistance to autoproteolysis results in greater
 CC proteolytic activity. The sequences in AAW62231 to AAW62265 represent
 CC subtilase fragments from the present invention
 XX
 XX Sequence 60 AA;
 SQ

Query Match 100.0%; Score 29; DB 2; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GPGS 5
 Db 28 GPGS 32

RESULT 39
 AAU48941
 ID AAU48941 standard; protein; 61 AA.
 XX
 XX AAU48941;
 AC
 XX
 XX 27-FEB-2002 (first entry)
 DT
 XX
 XX Propionibacterium acnes immunogenic protein #9837.
 DE
 XX
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 XX Propionibacterium acnes.
 OS
 XX WO200181581-A2.
 PN
 XX
 XX 01-NOV-2001.
 PD
 XX
 XX 20-APR-2001; 2001WO-US012865.
 PF
 XX
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59543.
 DR
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 XX Example 1; SEQ ID NO 10136; 1069pp; English.
 PS
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 61 AA;
 SQ

Query Match 100.0%; Score 29; DB 4; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GPGS 5
 Db 50 GPGS 54

RESULT 40

ABM45460
ID ABM45460 standard; protein; 61 AA.

XX AC

XX AC

XX DT 20-OCT-2003 (first entry)

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #10136.

XX KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;

XX KW immunostimulant; immune response; vaccine.

XX OS Propionibacterium acnes.

XX PN WO2003033515-A1.

XX PD 24-APR-2003.

XX PF 11-OCT-2002; 2002WO-US032727.

XX PR 15-OCT-2001; 2001US-00978825.

XX PA (CORI-) CORIXA CORP.

XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX PI Barth B, Vallieve-Douglas J;

XX DR WPI; 2003-381789/36.

XX DR N-PSDB; ACF64472.

XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

XX PT or for stimulating an immune response specific for a P. acnes protein.

XX PS Example 1; SEQ ID NO 10136; 1481pp; English.

XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

XX CC encoding a Propionibacterium acnes protein. The invention also relates to

XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

XX CC immunogenic fragments of P. acnes polypeptides. The invention

XX CC additionally encompasses expression vectors and host cells comprising a

XX CC polynucleotide of the invention; antibodies against polypeptides of the

XX CC invention; fusion proteins comprising a polypeptide of the invention; a

XX CC method for stimulating an immune response specific for a P. acnes

XX CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,

XX CC via this method; a vaccine composition comprising P. acnes polypeptides,

XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or

XX CC antigen-presenting cells that express the polypeptide); a method and kit

XX CC for detecting or determining the presence or absence of P. acnes in a

XX CC patient; and a method for inhibiting the development of P. acnes in a

XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

XX CC proteins, T cell populations or antigen-presenting cells that express the

XX CC polypeptides are useful for diagnosing, preventing or treating acne

XX CC vulgaris, or for stimulating an immune response specific for a P. acnes

XX CC protein. The polynucleotides can also be used as probes or primers for

XX CC nucleic acid hybridisation. The vaccine composition is useful for the

XX CC stimulation of an immune response against P. acnes, or for treating acne,

XX CC and the kit is useful for performing a diagnostic assay. The present

XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open

XX CC reading frame) contained within the P. acnes polynucleotides of the

XX CC invention. Note: The sequence data for this patent did not form part of

XX CC the printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 61 AA;

Query Match

Best Local Similarity 100.0%; Score 29; DB 6; Length 61;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:19:44 ; Search time 158 Seconds
(without alignments)

12.392 Million cell updates/sec

Title: TUNGAP-CLAIM8

Perfect score: 29

Sequence: 1 gpggs 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	8	9	US-09-910-071-13
2	29	100.0	8	16	US-10-479-638-40
3	29	100.0	10	9	US-09-815-837-111
4	29	100.0	10	10	US-09-572-404B-429
5	29	100.0	10	16	US-10-468-655-2
6	29	100.0	14	9	US-09-815-837-109
7	29	100.0	29	17	US-10-362-332-4
8	29	100.0	35	14	US-10-351-641-636
9	29	100.0	35	16	US-10-425-115-202567
10	29	100.0	39	14	US-10-029-386-28313
11	29	100.0	44	15	US-10-424-599-159438
Sequence 13, Appl					
Sequence 40, Appl					
Sequence 111, App					
Sequence 429, App					
Sequence 2, Appli					
Sequence 109, App					
Sequence 4, Appli					
Sequence 636, App					
Sequence 202567,					
Sequence 28313, A					
Sequence 159438,					

29	100.0	47	16	US-10-479-638-43	Sequence 43, Appl
29	100.0	51	11	US-09-864-408A-7476	Sequence 7476, Ap
29	100.0	53	15	US-10-424-599-232433	Sequence 232433,
29	100.0	54	16	US-10-767-701-50865	Sequence 50865, A
29	100.0	56	9	US-09-864-761-47371	Sequence 47371, A
29	100.0	58	9	US-09-948-080-13	Sequence 13, Appl
29	100.0	58	16	US-10-425-115-192067	Sequence 192067,
29	100.0	59	11	US-09-864-408A-8312	Sequence 8312, Ap
29	100.0	65	14	US-10-029-386-32946	Sequence 32946, A
29	100.0	65	16	US-10-767-701-51268	Sequence 51268, A
29	100.0	66	11	US-09-864-408A-3852	Sequence 3852, Ap
29	100.0	68	16	US-10-425-115-329137	Sequence 329137,
29	100.0	70	16	US-10-425-115-253228	Sequence 253228,
29	100.0	72	16	US-10-425-115-368339	Sequence 368339,
29	100.0	74	15	US-10-424-599-265256	Sequence 265256,
29	100.0	75	16	US-10-425-115-241011	Sequence 241011,
29	100.0	76	14	US-10-029-386-28551	Sequence 28551, A
29	100.0	77	15	US-10-424-599-228862	Sequence 228862,
29	100.0	78	9	US-09-886-426-4	Sequence 4, Appli
29	100.0	83	16	US-10-767-701-57393	Sequence 57393, A
29	100.0	83	16	US-10-425-115-248851	Sequence 248851,
29	100.0	84	15	US-10-424-599-198068	Sequence 198068,
29	100.0	84	16	US-10-425-115-318850	Sequence 318850,
29	100.0	86	15	US-10-424-599-182267	Sequence 182267,
29	100.0	86	16	US-10-437-963-159102	Sequence 159102,
29	100.0	86	16	US-10-767-701-47539	Sequence 47539, A
29	100.0	87	9	US-09-995-494-99	Sequence 99, Appl
29	100.0	87	16	US-10-425-115-236310	Sequence 236310,
29	100.0	88	14	US-10-029-386-23012	Sequence 23012, A
29	100.0	88	16	US-10-425-115-318598	Sequence 318598,
29	100.0	88	16	US-10-617-316-140	Sequence 140, App
29	100.0	90	16	US-10-767-701-54507	Sequence 54507, A
29	100.0	90	16	US-10-425-115-220563	Sequence 220563,
29	100.0	91	15	US-10-424-599-223094	Sequence 223094,
29	100.0	91	16	US-10-437-963-122587	Sequence 122587,
29	100.0	92	16	US-10-437-963-166924	Sequence 166924,
29	100.0	93	16	US-10-425-115-225472	Sequence 225472,
29	100.0	94	15	US-10-424-599-161713	Sequence 161713,
29	100.0	95	16	US-10-437-963-196805	Sequence 196805,
29	100.0	95	16	US-10-425-115-238567	Sequence 238567,
29	100.0	96	9	US-09-764-884-34	Sequence 34, Appl
29	100.0	96	14	US-10-092-256-34	Sequence 34, Appl
29	100.0	97	10	US-09-734-017A-40	Sequence 40, Appl
29	100.0	97	10	US-09-764-891-4970	Sequence 4970, Ap
29	100.0	98	16	US-10-437-963-125633	Sequence 125633,
29	100.0	98	16	US-10-425-115-257703	Sequence 257703,
29	100.0	99	15	US-10-424-599-279835	Sequence 279835,
29	100.0	102	15	US-10-425-114-49680	Sequence 49680, A
29	100.0	103	16	US-10-437-963-145205	Sequence 145205,
29	100.0	103	16	US-10-437-963-148057	Sequence 148057,
29	100.0	103	16	US-10-425-115-320034	Sequence 320034,
29	100.0	105	16	US-10-767-701-61893	Sequence 61893, A
29	100.0	107	16	US-10-425-115-205265	Sequence 205265,
29	100.0	107	16	US-10-425-115-270088	Sequence 270088,
29	100.0	110	16	US-10-767-701-33748	Sequence 33748, A
29	100.0	111	16	US-10-425-115-286070	Sequence 286070,
29	100.0	112	15	US-10-424-599-269136	Sequence 269136,
29	100.0	112	15	US-10-669-853-14	Sequence 14, Appl
29	100.0	112	16	US-10-661-984A-35	Sequence 35, Appl
29	100.0	112	18	US-10-356-264A-8	Sequence 8, Appli
29	100.0	113	9	US-09-320-920-3	Sequence 3, Appli
29	100.0	113	9	US-09-804-615-12	Sequence 12, Appl
29	100.0	113	15	US-10-424-599-190037	Sequence 190037,
29	100.0	113	15	US-10-669-853-13	Sequence 13, Appl
29	100.0	113	16	US-10-437-963-103822	Sequence 103822,
29	100.0	113	16	US-10-661-984A-12	Sequence 12, Appl
29	100.0	113	16	US-10-806-793-12	Sequence 12, Appl
29	100.0	113	16	US-10-451-567-1	Sequence 1, Appli
29	100.0	113	17	US-10-864-891-14	Sequence 14, Appl
29	100.0	113	18	US-10-356-264A-2	Sequence 2, Appli
29	100.0	113	18	US-10-356-264A-23	Sequence 23, Appl
29	100.0	114	9	US-09-804-615-37	Sequence 37, Appl
29	100.0	114	16	US-10-437-963-160846	Sequence 160846,

Sequence 54, Appl
Sequence 4, Appl
Sequence 11, Appl
Sequence 12, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 13, Appl
Sequence 7, Appl
Sequence 31, Appl
Sequence 3475, Ap
Sequence 40094, A
Sequence 294849,
Sequence 147424,
Sequence 36, Appl
Sequence 59919, A
Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-910-071-13
; Sequence 13, Application US/09910071
; Patent No. US20020116146A1
; GENERAL INFORMATION:
; APPLICANT: Tomikawa, Mayumi
; APPLICANT: Aikawa, Seiichi
; APPLICANT: Matsuzawa, Fumiko
; TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Simil
; TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three
; TITLE OF INVENTION: Structures of Molecules
; FILE REFERENCE: 522.1921D2
; CURRENT APPLICATION NUMBER: US/09/910,071
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/014,867
; PRIOR FILING DATE: 1993-02-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 8
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: adenylate kinase (3ADK)
US-09-910-071-13

Query Match 100.0%; Score 29; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db 1 GPGGS 5

RESULT 2
US-10-479-638-40
; Sequence 40, Application US/10479638
; Publication No. US20040210956A1
; GENERAL INFORMATION:
; APPLICANT: Don A. Roth
; APPLICANT: Randolph V. Lewis
; APPLICANT: The University of Wyoming
; TITLE OF INVENTION: Expression of Spider Silk Proteins in Higher Plants
; FILE REFERENCE: WVO 02-0004US
; CURRENT APPLICATION NUMBER: US/10/479,638
; CURRENT FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: PCT/US02/18256
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/296,184
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Argiope
US-10-479-638-40

Query Match 100.0%; Score 29; DB 16; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db 3 GPGGS 7

RESULT 3
US-09-815-837-111
; Sequence 111, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Mediators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:downstream
; OTHER INFORMATION: linker for C0582 and C0589
US-09-815-837-111

Query Match 100.0%; Score 29; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db 6 GPGGS 10

RESULT 4
US-09-572-404B-429
; Sequence 429, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 429
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:

; OTHER INFORMATION: sequence located in ANX1 OR LPC1 at 24-33 and may interact with s
; OTHER INFORMATION: 430 in this patent.
US-09-572-404B-429

Query Match 100.0%; Score 29; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||||
Db 6 GGPGS 10

RESULT 5

US-10-468-655-2
; Sequence 2, Application US/1046855
; Publication No. US2004019784A1
; GENERAL INFORMATION:
; APPLICANT: Miano, Joseph M.
; APPLICANT: Streb, Jeffrey W.
; APPLICANT: Chen, Jiyan
; TITLE OF INVENTION: RETINOID INDUCIBLE PROTEINS OF VASCULAR SMOOTH MUSCLE
; FILE REFERENCE: 176/61022
; CURRENT APPLICATION NUMBER: US/10/468,655
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: 60/271,193
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/293,097
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: serine
; OTHER INFORMATION: carboxypeptidase substrate binding domain I of C.
; OTHER INFORMATION: elegans F2E12.1
US-10-468-655-2

Query Match 100.0%; Score 29; DB 16; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||||
Db 4 GGPGS 8

RESULT 6

US-09-815-837-109
; Sequence 109, Application US/09815837
; Patent No. US2002008241A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 109
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:downstream
; OTHER INFORMATION: linker for COS80 and COS87
US-09-815-837-109

Query Match 100.0%; Score 29; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||||
Db 10 GGPGS 14

RESULT 7

US-10-362-332-4
; Sequence 4, Application US/10362332
; Publication No. US2005010098A1
; GENERAL INFORMATION:
; APPLICANT: Prescott, Stephen
; APPLICANT: Pearman, Terrece
; APPLICANT: Castro-Faria-Neto, Hugo
; APPLICANT: Stafforini, Diana
; APPLICANT: McInyre, Thomas
; TITLE OF INVENTION: Human Cytidine Monophosphate (CMP) Kinase cDNA
; FILE REFERENCE: 1321.2.62
; CURRENT APPLICATION NUMBER: US/10/362,332
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/226,614
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-362-332-4

Query Match 100.0%; Score 29; DB 17; Length 29;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||||
Db 1 GGPGS 5

RESULT 8

US-10-351-641-636
; Sequence 636, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-636

Query Match 100.0%; Score 29; DB 14; Length 35;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
| | | | |
DB 14 GGPGS 18

RESULT 9

US-10-425-115-202567
; Sequence 202567, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 202567
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116331C.1.pep
US-10-425-115-202567

Query Match 100.0%; Score 29; DB 16; Length 35;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
| | | | |
DB 26 GGPGS 30

RESULT 10

US-10-029-386-28313
; Sequence 28313, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28313
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21 55.0
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.33
US-10-029-386-28313

Query Match 100.0%; Score 29; DB 14; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
| | | | |
DB 4 GGPGS 8

RESULT 11

US-10-424-599-159438
; Sequence 159438, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 159438
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_114993C.1.pep
US-10-424-599-159438

Query Match 100.0%; Score 29; DB 15; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
| | | | |
DB 34 GGPGS 38

RESULT 12

US-10-479-638-43
; Sequence 43, Application US/10479638
; Publication No. US20040210956A1
; GENERAL INFORMATION:
; APPLICANT: Don A. Roth
; APPLICANT: Randolph V. Lewis
; APPLICANT: The University of Wyoming
; TITLE OF INVENTION: Expression of Spider Silk Proteins in Higher Plants
; FILE REFERENCE: WYO.02-0004US
; CURRENT APPLICATION NUMBER: US/10/479, 638
; CURRENT FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: PCT/US02/18256
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/296,184
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic spider silk protein repeat
US-10-479-638-43

Query Match 100.0%; Score 29; DB 16; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5

```
Db          42 GGPGS 46
|||||
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50865
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3480-038-P1-K1-D4.pep
US-10-767-701-50865

Query Match      100.0%; Score 29; DB 16; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 GGPGS 5
|||||
Db          35 GGPGS 39
|||||

RESULT 16
US-09-864-761-47371
; Sequence 47371, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

RESULT 13
US-09-864-408A-7476
; Sequence 7476, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7476
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-7476

Query Match      100.0%; Score 29; DB 11; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 GGPGS 5
|||||
Db          25 GGPGS 29
|||||

RESULT 14
US-10-424-599-232433
; Sequence 232433, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 232433
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51910C.1.pep
US-10-424-599-232433

Query Match      100.0%; Score 29; DB 15; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 GGPGS 5
|||||
Db          11 GGPGS 15
|||||

RESULT 15
US-10-767-701-50865
; Sequence 50865, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47371
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC016057.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.44
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.46
; OTHER INFORMATION: SWISSPROT HIT: P57078, EVALUE 8.00e-04
; OTHER INFORMATION: EST_HUMAN HIT: AA972567.1, EVALUE 1.00e-27
US-09-864-761-47371
```

```
Query Match          100.0%; Score 29; DB 9; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GPGGS 5
        |||||
Db      26 GPGGS 30
```

RESULT 17

```
US-09-948-080-13
; Sequence 13, Application US/09948080
; Patent No. US20020102702A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: VAN DER OSTEN, CLAUS
```

```
APPLICANT: HALKIER, TORDEN
```

```
APPLICANT: ANDERSEN, CARSTEN
```

```
APPLICANT: BAUDITZ, PETER
```

```
APPLICANT: HANSEN, PETER KAMP
```

```
TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
```

```
FILE REFERENCE: 4946,200-US
```

```
CURRENT APPLICATION NUMBER: US/09/948,080
```

```
CURRENT FILING DATE: 2001-09-06
```

```
PRIOR APPLICATION NUMBER: US/08/963,851
```

```
PRIOR FILING DATE: 1997-11-04
```

```
NUMBER OF SEQ ID NOS: 35
```

```
SOFTWARE: FastSEQ for Windows Version 3.0
```

```
SEQ ID NO 13
```

```
LENGTH: 58
```

```
TYPE: PRT
```

```
ORGANISM: Bacillus subtilis
```

```
US-09-948-080-13
```

```
Query Match          100.0%; Score 29; DB 9; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GPGGS 5
        |||||
Db      27 GPGGS 31
```

RESULT 18

```
US-10-425-115-192067
; Sequence 192067, Application US/10425115
; Publication No. US20040214272A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: La Rosa, Thomas J.
```

```
APPLICANT: Kovalic, David K.
```

```
APPLICANT: Zhou, Yihua
```

```
APPLICANT: Cao, Yongwei
```

```
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
```

```
TITLE OF INVENTION: Plants
```

```
FILE REFERENCE: 38-21(53222) B
```

```
CURRENT APPLICATION NUMBER: US/10/425,115
```

```
CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 369336
; SEQ ID NO 192067
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_106750C.1.pap
US-10-425-115-192067
```

```
Query Match          100.0%; Score 29; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GPGGS 5
        |||||
Db      1 GPGGS 5
```

RESULT 19

```
US-09-864-408A-8312
; Sequence 8312, Application US/09864408A
; Publication No. US20040009474A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: Leach, Martin D.
```

```
APPLICANT: Shimkets, Richard A.
```

```
TITLE OF INVENTION: No US20040009474A1el Human Polynucleotides and Polypeptides Encod
```

```
FILE REFERENCE: 21402-012
```

```
CURRENT APPLICATION NUMBER: US/09/864,408A
```

```
CURRENT FILING DATE: 2001-05-24
```

```
PRIOR APPLICATION NUMBER: 60/206,690
```

```
PRIOR FILING DATE: 2000-05-24
```

```
NUMBER OF SEQ ID NOS: 9068
```

```
SOFTWARE: FastSEQ for Windows Version 4.0
```

```
SEQ ID NO 8312
```

```
LENGTH: 59
```

```
TYPE: PRT
```

```
ORGANISM: Homo sapiens
```

```
US-09-864-408A-8312
```

```
Query Match          100.0%; Score 29; DB 11; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GPGGS 5
        |||||
Db      51 GPGGS 55
```

RESULT 20

```
US-10-029-386-32946
```

```
; Sequence 32946, Application US/10029386
```

```
; Publication No. US20030194704A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: Penn, Sharron G.
```

```
APPLICANT: Rank, David K.
```

```
APPLICANT: Hanzel, David K.
```

```
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
```

```
FILE REFERENCE: AEOMICA-X-2
```

```
CURRENT APPLICATION NUMBER: US/10/029,386
```

```
CURRENT FILING DATE: 2001-12-20
```

```
NUMBER OF SEQ ID NOS: 34288
```

```
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
```

```
SEQ ID NO 32946
```

```
LENGTH: 65
```

```
TYPE: PRT
```

```
ORGANISM: Homo sapiens
```

```
FEATURE:
```

```
OTHER INFORMATION: MAP TO U34879.1
```

```
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
```

```
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
```

```
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
```

```
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
```

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: SWISSPROT HIT: P17656, EVALUE 2.00e-03
US-10-029-386-32946

Query Match 100.0%; Score 29; DB 14; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||
Db 51 GGPGS 55

RESULT 21

US-10-767-701-51268
; Sequence 51268, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 51268
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB5047-011-R1-XP1-E12.pep
US-10-767-701-51268

Query Match 100.0%; Score 29; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||
Db 1 GGPGS 5

RESULT 22

US-09-864-408A-3852
; Sequence 3852, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3852
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-3852

Query Match 100.0%; Score 29; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||
Db 13 GGPGS 17

RESULT 23

US-10-425-115-329137
; Sequence 329137, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 329137
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_63243C.1.pep
US-10-425-115-329137

Query Match 100.0%; Score 29; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||
Db 58 GGPGS 62

RESULT 24

US-10-425-115-253228
; Sequence 253228, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253228
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162524C.1.pep
US-10-425-115-253228

Query Match 100.0%; Score 29; DB 16; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|||
Db 14 GGPGS 18

RESULT 25

US-10-425-115-368339
; Sequence 368339, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 369339
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99097C.1.pep
US-10-425-115-368339

Query Match 100.0%; Score 29; DB 16; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 30 GGPGS 34

RESULT 26
US-10-424-599-265256
; Sequence 265256, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265256
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81545C.1.pep
US-10-424-599-265256

Query Match 100.0%; Score 29; DB 15; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 68 GGPGS 72

RESULT 27
US-10-425-115-241011
; Sequence 241011, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 241011

; LENGTH: 75
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_151380C.1.pep
US-10-425-115-241011

Query Match 100.0%; Score 29; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 69 GGPGS 73

RESULT 28
US-10-029-386-29551
; Sequence 29551, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29551
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
US-10-029-386-29551

Query Match 100.0%; Score 29; DB 14; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 46 GGPGS 50

RESULT 29
US-10-424-599-228862
; Sequence 228862, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228862
; LENGTH: 77

```
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4868C.1.pep
US-10-424-599-228862

Query Match      100.0%; Score 29; DB 15; Length 77;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
      |||||
Db      5 GGPGS 9

RESULT 30
US-09-886-426-4
; Sequence 4, Application US/09886426
; Patent No. US20020128455A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn AB
; TITLE OF INVENTION: Growth Factors
; FILE REFERENCE: 00295
; CURRENT APPLICATION NUMBER: US/09/886,426
; CURRENT FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-426-4

Query Match      100.0%; Score 29; DB 9; Length 78;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
      |||||
Db      16 GGPGS 20

RESULT 31
US-10-767-701-57393
; Sequence 57393, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 57393
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30966640.pep
US-10-767-701-57393

Query Match      100.0%; Score 29; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
      |||||
Db      13 GGPGS 17

RESULT 32
US-10-425-115-248851
; Sequence 248851, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 248851
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(83)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_15853C.1.pep
US-10-425-115-248851

Query Match      100.0%; Score 29; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
      |||||
Db      30 GGPGS 34

RESULT 33
US-10-424-599-198068
; Sequence 198068, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 198068
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20880C.1.pep
US-10-424-599-198068

Query Match      100.0%; Score 29; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
      |||||
Db      28 GGPGS 32

RESULT 34
```

```
US-10-425-115-318850
; Sequence 318850, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 318850
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_53865C.1.pep
US-10-425-115-318850

Query Match      100.0%; Score 29; DB 16; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GPGGS 5
Db      2 GPGGS 6

RESULT 35
US-10-424-599-182267
; Sequence 182267, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 182267
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_1355C.1.pep
US-10-424-599-182267

Query Match      100.0%; Score 29; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GPGGS 5
Db      34 GPGGS 38

RESULT 36
US-10-437-963-159102
; Sequence 159102, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

US-10-425-115-318850
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159102
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58511C.1.pep
US-10-437-963-159102

Query Match      100.0%; Score 29; DB 16; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GPGGS 5
Db      82 GPGGS 86

RESULT 37
US-10-767-701-47539
; Sequence 47539, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 47539
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3481-044-Q6-K1-H12.pep
US-10-767-701-47539

Query Match      100.0%; Score 29; DB 16; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GPGGS 5
Db      39 GPGGS 43

RESULT 38
US-09-995-494-99
; Sequence 99, Application US/09995494
; Patent No. US20020127578A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
```



```
; APPLICANT: Chen, Sei-Yu
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0293
; CURRENT APPLICATION NUMBER: US/09/995,494
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/253,176
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-995-494-99
```

```
Query Match 100.0%; Score 29; DB 9; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GPGS 5
    |||||
Db 28 GPGS 32
```

RESULT 39

```
US-10-425-115-236310
; Sequence 236310, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 236310
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_147097C.1.pep
US-10-425-115-236310
```

```
Query Match 100.0%; Score 29; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GPGS 5
    |||||
Db 50 GPGS 54
```

RESULT 40

```
US-10-029-386-29012
; Sequence 29012, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEOWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29012
; LENGTH: 88
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: SWISSPROT HIT: Q9WUAI, EVALUAE 1.00e-03
US-10-029-386-29012
```

```
Query Match 100.0%; Score 29; DB 14; Length 88;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GPGS 5
    |||||
Db 30 GPGS 34
```

Search completed: August 19, 2005, 16:29:37
Job time : 160 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:16:13 ; Search time 39 Seconds

(without alignments)
12.335 Million cell updates/sec

Title: TUNGAP-CLAIMS

Perfect score: 29

Sequence: 1 gppgs 5

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 79:*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	30	A26188	lipocortin I - pig
2	29	100.0	64	T05913	hypothetical prote
3	29	100.0	70	S20770	Ig heavy chain V r
4	29	100.0	102	T17892	hypothetical prote
5	29	100.0	112	T22078	hypothetical prote
6	29	100.0	114	F72484	hypothetical prote
7	29	100.0	131	T34900	probable secreted
8	29	100.0	135	B38641	inositol 1,4,5-tri
9	29	100.0	136	E72759	hypothetical prote
10	29	100.0	143	JH0385	midkine precursor
11	29	100.0	146	T02395	hypothetical prote
12	29	100.0	152	F86473	hypothetical prote
13	29	100.0	154	H69063	conserved hypothet
14	29	100.0	167	B96733	hypothetical prote
15	29	100.0	175	A38641	inositol 1,4,5-tri
16	29	100.0	180	T28938	hypothetical prote
17	29	100.0	181	S74491	adenylate kinase (
18	29	100.0	182	F72623	hypothetical prote
19	29	100.0	193	KICAC	adenylate kinase (
20	29	100.0	193	T00585	hypothetical prote
21	29	100.0	194	KIBOA	adenylate kinase (
22	29	100.0	194	KIHUA	adenylate kinase (
23	29	100.0	194	KIPGA	adenylate kinase (
24	29	100.0	194	KIRBA	adenylate kinase (
25	29	100.0	194	A25327	adenylate kinase (
26	29	100.0	194	A35235	adenylate kinase (
27	29	100.0	200	G72471	hypothetical prote
28	29	100.0	210	T21947	hypothetical prote
29	29	100.0	210	T47842	URIDYLATE KINASE-1

30	29	100.0	212	JC7511	fibroblast growth
31	29	100.0	219	A81611	ribosomal protein
32	29	100.0	222	A88102	protein W09G10.1 [
33	29	100.0	225	A72056	L3 ribosomal prote
34	29	100.0	225	D86571	L3 ribosomal prote
35	29	100.0	225	C87595	isochorismatase fa
36	29	100.0	227	T05549	adenylate kinase h
37	29	100.0	230	I37095	gene 2.19 protein
38	29	100.0	235	T50170	hypothetical prote
39	29	100.0	237	A48912	leucine zipper pro
40	29	100.0	239	T01599	hypothetical prote
41	29	100.0	244	A64843	hypothetical prote
42	29	100.0	244	A90786	probable synthetas
43	29	100.0	244	G85645	probable synthetas
44	29	100.0	246	AH2883	isochorismatase li
45	29	100.0	246	F97859	probable synthetas
46	29	100.0	267	B84089	hypothetical prote
47	29	100.0	269	T04394	NBS-LRR type resis
48	29	100.0	287	A47038	prolyl aminopeptid
49	29	100.0	291	T31592	hypothetical prote
50	29	100.0	294	D72395	sugar ABC transpor
51	29	100.0	294	A59087	prolyl aminopeptid
52	29	100.0	299	T43442	hypothetical prote
53	29	100.0	305	T36056	hypothetical prote
54	29	100.0	305	T30165	hypothetical prote
55	29	100.0	307	A45581	Distal-less homeob
56	29	100.0	308	B64202	prolyl aminopeptid
57	29	100.0	309	YAZQN7	S-antigen precursor
58	29	100.0	309	S73458	prolyl aminopeptid
59	29	100.0	313	B82671	proline imino-pept
60	29	100.0	316	C96733	hypothetical prote
61	29	100.0	322	AD3044	hypothetical prote
62	29	100.0	322	H98241	ABC transporter su
63	29	100.0	323	T35734	probable aminopept
64	29	100.0	323	B83010	prolyl aminopeptid
65	29	100.0	323	T24582	hypothetical prote
66	29	100.0	328	G02469	homeotic protein D
67	29	100.0	332	JH0465	homeotic protein T
68	29	100.0	333	T23618	hypothetical prote
69	29	100.0	335	A96733	hypothetical prote
70	29	100.0	345	A48462	dense granule prot
71	29	100.0	346	LUGP1	annexin I - guinea
72	29	100.0	346	LUDU	annexin I - human
73	29	100.0	346	LUMS1	annexin I - mouse
74	29	100.0	346	LURT1	annexin I - rat
75	29	100.0	346	S28228	annexin I - bovine
76	29	100.0	347	T06584	probable DNA-bind
77	29	100.0	349	T67418	transcription fact
78	29	100.0	349	T50369	transcription fact
79	29	100.0	349	T51739	transcription fact
80	29	100.0	349	T53277	transcription fact
81	29	100.0	349	T67417	transcription fact
82	29	100.0	353	T09665	peroxidase (EC 1.1
83	29	100.0	358	T26281	hypothetical prote
84	29	100.0	359	B64919	conserved hypothet
85	29	100.0	359	D90920	hypothetical prote
86	29	100.0	359	A85769	hypothetical prote
87	29	100.0	362	A70547	probable menE - My
88	29	100.0	364	T24153	hypothetical prote
89	29	100.0	368	T24153	hypothetical prote
90	29	100.0	371	AH0691	o-succinylbenzoic
91	29	100.0	375	TJ0427	probable oxidoredu
92	29	100.0	375	TJ0427	S-antigen precursor
93	29	100.0	376	D87451	oxidoreductase, Gf
94	29	100.0	392	G87233	probable lipoprote
95	29	100.0	396	T01201	hypothetical prote
96	29	100.0	404	B84745	probable RNA-bind
97	29	100.0	410	T36535	probable hydrolase
98	29	100.0	414	T48975	apolipoprotein A-I
99	29	100.0	414	A37133	apolipoprotein A-I
100	29	100.0	416	D83386	hypothetical prote
100	29	100.0	419	T15088	hypothetical prote

ALIGNMENTS

```
RESULT 1
A26188
lipocortin I - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 23-Feb-1997
C;Accession: A26188
J;De, B.K.; Misono, K.S.; Lukas, T.J.; Mroczkowski, B.; Cohen, S.
R; Biol. Chem. 261, 13784-13792, 1986
A;Title: A calcium-dependent 35-kilodalton substrate for epidermal growth factor receptor
A;Reference number: A26188; MUID:87008618; PMID:3020049
A;Accession: A26188
A;Molecule type: protein
A;Residues: 1-30 <DEB>
C;Superfamily: annexin I; annexin repeat homology
C;Keywords: phosphoprotein

Query Match 100.0%; Score 29; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 18 GPGGS 22

RESULT 2
T05913
hypothetical protein - barley (fragment)
C;Species: Hordeum vulgare (barley)
C;Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05913
R;Hess, W.R.; Golz, R.R.; Boerner, T.
Plant Sci. 133, 191-201, 1998
A;Title: Analysis of randomly selected cDNAs reveals the expression of stress- and defense
A;Reference number: Z15411
A;Accession: T05913
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-64 <HES>
A;Cross-references: UNIPROT:O23989; EMBL:AJ000232; NID:e1055248; PIDN:CAA03958.1; PID:el
A;Experimental source: cv. Hailea, leaf

Query Match 100.0%; Score 29; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 13 GPGGS 17

RESULT 3
S20770
Ig heavy chain V region - human
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C;Accession: S20770
R;Mortari, F.; Wang, J.; Schroeder, H.W.
submitted to the EMBL Data Library, April 1992
A;Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.
A;Reference number: S20764
A;Accession: S20770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-70 <MOR>
A;Cross-references: EMBL:Z11953; NID:g33869; PIDN:CAA78010.1; PID:g33870
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 29; DB 2; Length 70;
```

```
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 44 GPGGS 48

RESULT 4
T17892
hypothetical protein a389R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17892
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17892
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-102 <GRA>
A;Cross-references: UNIPROT:Q98441; EMBL:U42580; NID:g4028896; PIDN:AAC96757.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: a389R
C;Superfamily: Chlorella virus PBCV-1 hypothetical protein a389R

Query Match 100.0%; Score 29; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 52 GPGGS 56

RESULT 5
T22078
hypothetical protein F41E7.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22078
R;Lennard, N.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19509
A;Accession: T22078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-112 <WIL>
A;Cross-references: UNIPROT:Q20277; EMBL:Z68106; PIDN:CAA92128.1; GSPDB:GNO0028; CESP:F41
A;Experimental source: clone F41E7
C;Genetics:
A;Gene: CESP:F41E7.5
A;Map position: X
C;Superfamily: Arabidopsis glycine-rich protein 3

Query Match 100.0%; Score 29; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 51 GPGGS 55

RESULT 6
F72484
hypothetical protein APE2518 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72484
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
```

DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <KAW>
A;Cross-references: UNIPROT:Q9Y8W6; DDBJ:AP000064; NID:g5105945; PIDN:BAA81534.1; PID:dl
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2518
C;Superfamily: Aeropyrum pernix hypothetical protein APE2518

Query Match 100.0%; Score 29; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
|||||
Db 39 GPGGS 43

RESULT 7
T34900
Probable secreted protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34900
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21561
A;Accession: T34900
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-131 <OLI>
A;Cross-references: UNIPROT:O69934; EMBL:AL023861; PIDN:CAA19617.1; GSPDB:GN00070; SCOE
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOE:DB:SC3C8.25C
C;Superfamily: Streptomyces coelicolor probable secreted protein SC3C8.25C

Query Match 100.0%; Score 29; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
|||||
Db 107 GPGGS 111

RESULT 8
B38641
Inositol 1,4,5-trisphosphate receptor, short splice form - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 18-Jun-1993
C;Accession: B38641
R;Danoff, S.K.; Ferris, C.D.; Donath, C.; Fischer, G.A.; Munemitsu, S.; Ullrich, A.; Sny
Proc. Natl. Acad. Sci. U.S.A. 88, 2951-2955, 1991
A;Title: Inositol 1,4,5-trisphosphate receptors: distinct neuronal and nonneuronal forms
A;Reference number: A38641; MUID:91187909; PMID:1849282
A;Accession: B38641
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-135 <DAN>
A;Cross-references: GB:M64698
C;Superfamily: inositol-trisphosphate receptor

Query Match 100.0%; Score 29; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
|||||

Db 88 GPGGS 92

RESULT 9

E72759
Hypothetical protein APE0066 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E72759
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-136 <KAW>
A;Cross-references: UNIPROT:Q9YG35; DDBJ:AP000058; NID:g5103388; PIDN:BAA78975.1; PID:dl
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0066
C;Superfamily: Aeropyrum pernix hypothetical protein APE0066

Query Match 100.0%; Score 29; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
|||||
Db 115 GPGGS 119

RESULT 10

JH0385
Midkine precursor - human
N;Alternate names: amphiregulin-associated protein, large form; neurite outgrowth-promot
N;Contains: amphiregulin-associated protein, short form
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: JH0385; A38798; S14467; JQ1196; I53002; I55196
R;Tsubui, J.; Uehara, K.; Kadomatsu, K.; Matsubara, S.; Muramatsu, T.
Biochem. Biophys. Res. Commun. 176, 792-797, 1991
A;Title: A new family of heparin-binding factors: strong conservation of midkine (MK) seq
A;Reference number: JH0385; MUID:91222250; PMID:2025291
A;Accession: JH0385
A;Molecule type: mRNA
A;Residues: 26-143 <TS1>
A;Cross-references: UNIPROT:P21741
A;Accession: A38798
A;Molecule type: DNA
A;Residues: 1-25 <TS2>
A;Experimental source: kidney
R;Kretschmer, P.J.; Fairhurst, J.L.; Decker, M.M.; Chan, C.P.; Gluzman, Y.; Bohlen, P.;
submitted to the EMBL Data Library, October 1990
A;Description: A human neurite outgrowth-promoting protein is a member of a novel gene f
A;Reference number: S14467
A;Accession: S14467
A;Molecule type: mRNA
A;Residues: 1-143 <KRE>
A;Cross-references: EMBL:X55110; NID:g35086; PIDN:CAA38908.1; PID:g35087
R;Shoyab, M.; McDonald, V.L.; Dick, K.; Modrell, B.; Malik, N.; Plozman, G.D.
Biochem. Biophys. Res. Commun. 179, 572-578, 1991
A;Title: Amphiregulin-associated protein: complete amino acid sequence of a protein prod
A;Reference number: JQ1196; MUID:91354311; PMID:1883381
A;Accession: JQ1196
A;Molecule type: protein
A;Residues: 21-143 <SHO>
R;Fairhurst, J.L.; Kretschmer, P.J.; Kovacs, E.; Bohlen, P.; Kovsed, I.
DNA Cell Biol. 12, 139-147, 1993
A;Title: Structure of the gene coding for the human retinoic acid-inducible factor, MK.
A;Reference number: I53002; MUID:93228828; PMID:8471163
A;Accession: I53002

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-143 <RES>
A;Cross-references: GB:M94250; NID:g188570; PIDN:AAA59850.1; PID:g188571
R;Uehara, K.; Matsubara, S.; Kadomatsu, K.; Tsutsui, J.; Muramatsu, T.
J. Biochem. 111, 563-567, 1992
A;Title: Genomic structure of human midline (MK), a retinoic acid-responsive growth/differentiation factor
A;Reference number: 155196; MUID:92348340; PMID:1639750
A;Accession: 155196
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-143 <RES>
A;Cross-references: GB:D10604; NID:g219928; PIDN:BAA01457.1; PID:g219929
C;Genetics:
A;Gene: GDB:MDK; NEG2
A;Cross-references: GDB:143094; OMIM:162096
A;Map position: 11p11.2-11p11.2
A;Introns: 26/1; 82/1; 136/1
C;Superfamily: pleiotrophin
C;Keywords: growth factor; heparin binding
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-143/Product: midline, long form #status experimental <MA2>
F;23-143/Product: midline, short form #status experimental <MA2>

Query Match 100.0%; Score 29; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGS 5
Db 31 GPGS 35

RESULT 11
T02395
hypothetical protein At2g44400 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F411.21
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
A;Accession: T02395; A84878
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84878
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <ROU>
A;Cross-references: UNIPROT:O64874; EMBL:AC004521; NID:g3128166; PID:g3128182
A;Experimental source: cultivar Columbia
M.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84878
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <STO>
A;Cross-references: GB:AE002093; NID:g3128182; PIDN:AAC16086.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g44400; F411.21
A;Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGS 5
Db 109 GPGS 113

RESULT 12

F86473

hypothetical protein T911.17 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: F86473

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F86473

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-152 <STO>

A;Cross-references: UNIPROT:Q9C717; GB:AE005172; NID:g11034944; PIDN:AAG27101.1; GSPDB:G

C;Genetics:

C;Superfamily: wound-induced protein Sn-1

A;Map position: 1

Query Match 100.0%; Score 29; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 GPGS 5
Db 108 GPGS 112

RESULT 13

H69063

conserved hypothetical protein MTH1478 - Methanobacterium thermoautotrophicum (strain Del

C;Species: Methanobacterium thermoautotrophicum

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: H69063

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.;

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: H69063

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-154 <MTH>

A;Cross-references: UNIPROT:O27522; GB:AE000908; GB:AE000666; NID:g2622579; PIDN:AA8595;

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1478

C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1587

Query Match 100.0%; Score 29; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 GPGS 5
Db 141 GPGS 145

RESULT 14

B96733

hypothetical protein F15H11.9 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: B96733

R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B96733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-167 <STO>
A;Cross-references: UNIPROT:Q8RXL9; GB:AE005173; NID:95902397; PIDN:AAD55499.1; GSPDB:GN
C;Genetics:
A;Gene: F15H11.9
A;Map position: 1
C;Superfamily: wound-induced protein Sn-1

Query Match 100.0%; Score 29; DB 2; Length 167;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPGS 5
Db 123 GPGS 127

RESULT 15
A38641
inositol 1,4,5-trisphosphate receptor, long splice form - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 18-Jun-1993
C;Accession: A38641
R;Danoff, S.K.; Ferris, C.D.; Donath, C.; Fischer, G.A.; Munemitsu, S.; Ullrich, A.; Sny Proc. Natl. Acad. Sci. U.S.A. 88, 2951-2955, 1991
A;Title: Inositol 1,4,5-trisphosphate receptors: distinct neuronal and nonneuronal forms
A;Reference number: A38641; MUID:91187909; PMID:1849282
A;Accession: A38641
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-175 <DAN>
A;Cross-references: GB:M64699
C;Superfamily: inositol-trisphosphate receptor

Query Match 100.0%; Score 29; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPGS 5
Db 128 GPGS 132

RESULT 16
T28938
hypothetical protein F07C4.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28938
R;Miller, N.; Stellyes, L.
submitted to the EMBL Data Library, January 1997
A;Description: The sequence of C. elegans cosmid F07C4.
A;Reference number: Z20546
A;Accession: T28938
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-180 <MIL>
A;Cross-references: UNIPROT:P91207; EMBL:U80023; PIDN:AAC48014.1; GSPDB:GN000023; CESP:FO
A;Experimental source: strain Bristol N2; clone F07C4
C;Genetics:

A;Gene: CBSP:F07C4.7
A;Map position: 5
A;Introns: 58/1; 71/1; 86/1; 107/1; 127/1; 147/1; 165/1
C;Superfamily: glycine-rich cell wall structural protein 1

Query Match 100.0%; Score 29; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 30 GPGS 34

RESULT 17

S74491
adenylate kinase (EC 2.7.4.3) 1 - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein sl11059
C;Species: Synechocystis sp.

A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74491
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; o, K.; Okumura, T.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

8.
A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S74491
A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
A;Residues: 1-181 <KAN>

A;Cross-references: UNIPROT:P72641; EMBL:D90899; GB:AB001339; NID:G1651650; PIDN:BAAL1664;
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Genetics:

A;Gene: adk-1
C;Superfamily: adenylate kinase

C;Keywords: ATP; phosphotransferase

Query Match 100.0%; Score 29; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 7 GPGS 11

RESULT 18

F72623
hypothetical protein APE1446 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: F72623

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: F72623

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-182 <KAW>

A;Cross-references: UNIPROT:Q9YC04; DBJ:AP000061; NID:G5104821; PIDN:BAAS0444.1; PID:d1

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE1446

Query Match 100.0%; Score 29; DB 2; Length 182;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5

```

Db          142 GPGS 146
|||||
RESULT 19
KICAC
adenylate kinase (EC 2.7.4.3), cytosolic - common carp
C;Species: Cyprinus carpio (common carp)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: S00394; A31220
R;Reuner, C.; Hable, M.; Wilmanns, M.; Kiefer, E.; Schiltz, E.; Schulz, G.E.
Protein Seq. Data Anal. 1, 335-343, 1988
A;Title: Amino acid sequence and three-dimensional structure of cytosolic adenylate kinase
A;Reference number: S00394; MUID:89128814; PMID:2851785
A;Accession: S00394
A;Molecule type: protein
A;Residues: 1-193 <REU>
A;Cross-references: UNIPROT:P12115
C;Function:
A;Description: catalyzes the reversible phosphorylation of adenine monophosphate with ad
C;Superfamily: adenylate kinase
C;Keywords: acetylated amino end; ATP; muscle; P-loop; phosphotransferase
F;14-21/Region: nucleotide-binding motif A (P-loop) #status atypical
F;88-93/Region: nucleotide-binding motif B #status atypical
F;1/Modified site: acetylated amino end (Ala) #status experimental
F;24,35,37,92/Active site: Cys, His, Ser, Asp #status predicted
F;43,127,137,148/Binding site: substrate phosphate (Arg) #status predicted
Query Match      100.0%; Score 29; DB 1; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 GPGS 5
|||||
Db          14 GPGS 18
|||||
RESULT 20
T00585
hypothetical protein At2g30130 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein T27E13.13
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C;Accession: T00585; H84704
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
Submitted to the EMBL Data Library, May 1998
A;Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.
A;Reference number: Z14178
A;Accession: T00585
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-193 <ROU>
A;Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150407
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84704
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-193 <STO>
A;Cross-references: GB:AE002093; NID:g3150407; PIDN:AAC16959.1; GSPDB:GN00139
C;Genetics:
A;Gene: T27E13.13; At2g30130
A;Map position: 2
A;Introns: 54/3
C;Superfamily: Arabidopsis thaliana hypothetical protein T20E23.110

Query Match      100.0%; Score 29; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;

```

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 GPGS 5
|||||
Db          2 GPGS 6
|||||
RESULT 21
KIBOA
adenylate kinase (EC 2.7.4.3) - bovine
N;Alternate names: ATP-AMP transphosphorylase; myokinase
C;Species: Bos primigenius taurus (cattle)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C;Accession: A00681
R;Kuby, S.A.
unpublished results, cited by Kuby, S.A., Palmieri, R.H., Frischat, A., Fische, A.H., Wu,
A;Reference number: A00681
A;Accession: A00681
A;Molecule type: protein
A;Residues: 1-194 <KUB>
A;Cross-references: UNIPROT:P00570
C;Function:
A;Description: catalyzes the reversible phosphorylation of adenine monophosphate with ad
C;Superfamily: adenylate kinase
C;Keywords: acetylated amino end; ATP; P-loop; phosphotransferase
F;15-22/Region: nucleotide-binding motif A (P-loop) #status atypical
F;89-94/Region: nucleotide-binding motif B #status atypical
F;1/Modified site: acetylated amino end (Met) #status experimental
F;25,36,38,93/Active site: Cys, His, Ser, Asp #status predicted
F;44,128,138,149/Binding site: substrate phosphate (Arg) #status predicted
Query Match      100.0%; Score 29; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 GPGS 5
|||||
Db          15 GPGS 19
|||||
RESULT 22
KIHUA
adenylate kinase (EC 2.7.4.3) 1 - human (tentative sequence)
N;Alternate names: myokinase
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004
C;Accession: A33508; A00679
R;Matsuura, S.; Igarashi, M.; Tanizawa, Y.; Yamada, M.; Kishi, F.; Fujii, H.;
J. Biol. Chem. 264, 10148-10155, 1989
A;Title: Human adenylate kinase deficiency associated with hemolytic anemia. A single bas
A;Reference number: A33508; MUID:89255503; PMID:2542324
A;Accession: A33508
A;Molecule type: DNA
A;Residues: 1-126,'Q',128-180,'S',182-194 <MAT>
A;Cross-references: UNIPROT:P00568; GB:J04809; NID:g178321; PIDN:AAA51686.1; PID:g178322
A;Note: this sequence may be a less common allele found in Japanese populations
R;von Zabern, I.; Wittmann-Liebold, B.; Untucht-Grau, R.; Schirmer, R.H.; Pai, E.F.
Eur. J. Biochem. 68, 281-290, 1976
A;Title: Primary and tertiary structure of the principal human adenylate kinase.
A;Reference number: A00679; MUID:77003085; PMID:183954
A;Accession: A00679
A;Molecule type: protein
A;Residues: 1-194 <VON>
A;Experimental source: skeletal muscle
A;Note: this sequence represents the AK1 (1) allele, the most common of at least five alle
C;Comment: This form of the enzyme is expressed in the cytoplasm of skeletal muscle and i
C;Comment: A deficiency of this enzyme in erythrocytes is a cause of hemolytic anemia.
C;Genetics:
A;Gene: GDB:AK1
A;Cross-references: GDB:119664; OMIM:103000
A;Map position: 9q34.1-9q34.1
A;Introns: 3/1; 15/1; 69/3; 108/3; 172/3
A;Note: the first intron occurs before the initiator codon

```


C;Complex: monomer
C;Function:
A;Description: catalyzes the reversible phosphorylation of adenine monophosphate with ad
C;Superfamily: adenylate kinase
C;Keywords: acetylated amino end; ATP; P-loop; phosphotransferase
F;15-22/Region: nucleotide-binding motif A (P-loop) #status atypical
F;89-94/Region: nucleotide-binding motif B #status atypical
F;1/Modified site: acetylated amino end (Met) #status experimental
F;25,36,38,93/Active site: Cys, His, Ser, Asp #status predicted
F;44,128,138,149/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 15 GPGGS 19
|||||

RESULT 23
KIPGA
adenylate kinase (EC 2.7.4.3) [validated] - pig
N;Alternate names: myokinase
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 09-Jul-2004
C;Accession: A00682
R;Heil, A.; Mueller, G.; Noda, L.; Pinder, T.; Schirmer, H.; von Zabern, I.
Eur. J. Biochem. 43, 131-144, 1974
A;Title: The amino-acid sequence of porcine adenylate kinase from skeletal muscle.
A;Reference number: A91218; MUID:74267087; PMID:4366177
A;Accession: A00682
A;Molecule type: protein
A;Residues: 1-194 <HEI>
A;Cross-references: UNIPROT:P00571
A;Experimental source: skeletal muscle
R;Schulz, G.E.
submitted to the Brookhaven Protein Data Bank, November 1987
A;Reference number: A50563; PDB:3ADK
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 1-194
R;Dreusicke, D.; Karplus, P.A.; Schulz, G.E.
J. Mol. Biol. 199, 359-371, 1988
A;Title: Refined structure of porcine cytosolic adenylate kinase at 2.1 angstroms resolu
A;Reference number: A58438; MUID:88172480; PMID:2832612
A;Contents: annotation; X-ray crystallography, 2.1 angstroms
R;Schulz, G.E.; Elzinga, M.; Marx, F.; Schirmer, R.H.
Nature 250, 120-123, 1974
A;Title: Three-dimensional structure of adeny kinase.
A;Reference number: A93170; MUID:74277369; PMID:4367210
A;Contents: annotation; X-ray crystallography, 3.0 angstroms
C;Function:
A;Description: catalyzes the reversible phosphorylation of adenine monophosphate with ad
C;Superfamily: adenylate kinase
C;Keywords: acetylated amino end; ATP; P-loop; phosphotransferase; skeletal muscle
F;1-194/Product: adenylate kinase #status experimental <MAT>
F;15-22/Region: nucleotide-binding motif A (P-loop) #status atypical
F;89-94/Region: nucleotide-binding motif B #status atypical
F;1/Modified site: acetylated amino end (Met) #status experimental
F;25,36,38,93/Active site: Cys, His, Ser, Asp #status predicted
F;44,128,138,149/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 15 GPGGS 19
|||||

RESULT 24
KIRBA
adenylate kinase (EC 2.7.4.3) - rabbit

N;Alternate names: ATP-AMP transphosphorylase; myokinase
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C;Accession: A00680
R;Kuby, S.A.; Palmieri, R.H.; Frischat, A.; Fischer, A.H.; Wu, L.H.; Maland, L.; Manship,
Biochemistry 23, 2393-2399, 1984
A;Title: Studies on adenosine triphosphate transphosphorylases. Amino acid sequence of r
A;Reference number: A00680; MUID:85000428; PMID:6089869
A;Accession: A00680
A;Molecule type: protein
A;Residues: 1-194 <KUB>
A;Cross-references: UNIPROT:P00569
C;Function:
A;Description: catalyzes the reversible phosphorylation of adenine monophosphate with ad
C;Superfamily: adenylate kinase
C;Keywords: acetylated amino end; ATP; P-loop; phosphotransferase
F;15-22/Region: nucleotide-binding motif A (P-loop) #status atypical
F;89-94/Region: nucleotide-binding motif B #status atypical
F;1/Modified site: acetylated amino end (Met) #status experimental
F;25,36,38,93/Active site: Cys, His, Ser, Asp #status predicted
F;44,128,138,149/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 15 GPGGS 19
|||||

RESULT 25
adenylate kinase (EC 2.7.4.3) - chicken
N;Alternate names: ATP-AMP transphosphorylase; myokinase
C;Species: Gallus gallus (chicken)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A25327; A41458; A40243
R;Kishi, F.; Maruyama, M.; Tanizawa, Y.; Nakazawa, A.
J. Biol. Chem. 261, 2942-2945, 1986
A;Title: Isolation and characterization of cDNA for chicken muscle adenylate kinase.
A;Reference number: A25327; MUID:86140051; PMID:2419325
A;Accession: A25327
A;Molecule type: mRNA
A;Residues: 1-194 <KIS>
A;Cross-references: UNIPROT:P05081; GB:M12153; NID:g211125; PIDN:AAB59961.1; PID:g211126
R;Suminami, Y.; Kishi, F.; Torigoe, T.; Nakazawa, A.
J. Biochem. 103, 611-617, 1988
A;Title: Structure and complete nucleotide sequence of the gene encoding chicken cytosol
A;Reference number: A41458; MUID:89008161; PMID:2844738
A;Accession: A41458
A;Molecule type: DNA
A;Residues: 1-194 <SUM>
A;Cross-references: GB:D00251; NID:g222785; PIDN:BAA00182.1; PID:g222786
R;Salvucci, M.E.; Chavan, A.J.; Haley, B.E.
Biochemistry 31, 4479-4487, 1992
A;Title: Identification of peptides from the adenine binding domains of ATP and AMP in a
A;Reference number: A40243; MUID:92256377; PMID:1581304
A;Accession: A40243
A;Molecule type: protein
A;Residues: 28-36;125-146;154-177 <SAL>
A;Note: peptides corresponding to residues 28-44, 125-135, and 153-166 implicated in bin
R;Okajima, T.; Tanizawa, K.; Yoneya, T.; Fukui, T.
J. Biol. Chem. 266, 11442-11447, 1991
A;Title: Role of leucine 66 in the asymmetric recognition of substrates in chicken muscul
A;Reference number: A40419; MUID:91268001; PMID:2050660
A;Contents: annotation; site-directed mutagenesis
A;Note: residue Leu-67 implicated in binding site specificity for AMP rather than MgATP
C;Comment: This enzyme catalyzes the conversion of one molecule of AMP and one molecule
C;Genetics:
A;Introns: 4/1; 16/1; 70/3; 109/3; 173/3
C;Superfamily: adenylate kinase
C;Keywords: ATP; P-loop; phosphotransferase

F;16-23/Region: nucleotide-binding motif A (P-loop) #status atypical
F;90-95/Region: nucleotide-binding motif B #status atypical
F;26-37/Region: Active site: Cys, His, Ser, Asp #status predicted
F;45,129,139,150/Binding site: substrate phosphate (Arg #status pre-

Query Match	100.0%	Score 29;	DB 2;	Length 194;
Best Local Similarity	100.0%	Pred. No. 2.2e+02;		
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

Qy 1 GGPGS 5
db 16 GGPGS 20

RESULT 26

A35235
adenylate kinase (EC 2.7.4.3) - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 09-Jul-2004
C:Accession: A35235
J:Wiesmueller, L.; Noegel, A.A.; Barzu, O.; Gerisch, G.; Schleicher, M.
J. Biol. Chem. 265, 6339-6345, 1990
A:Title: cDNA-derived sequence of UMP-CMP kinase from Dictyostelium discoideum
A:Reference number: A35235; UID:90202912; PMID:2156849

Query Match	100.0%	Score 29;	DB 2;	Length 194;
Best Local Similarity	100.0%	Pred. No. 2.2e+02;		
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY 1 GGPGS 5
13 GGPGS 17
Db

RESULT 27

G72471
hypothetical protein APE2416 - Aeropyrum pernix (strain K1)
C/Species: Aeropyrum pernix
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: G72471
R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A/Reference number: A72450; MUID:99310339; PMID:10382966
A/Accession: G72471
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-200 <KAW>
A/Cross-references: UNIPROT:Q9Y969; DBJ:AP000064; NID:G5105945; PIDN:BAAB1431.1; PID:dl
A/Experimental source: strain K1
C/Genetics:
A/Gene: APE2416

Query Match	100.0%	Score 29;	DB 2;	Length 200;
Best Local Similarity	100.0%	Pred. No. 2.2e+02;		
Matches	5;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy 1 GGPGS 5
db 42 GGPGS 46

RESULT 28

T21947

hypothetical protein F38B2.4 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T21947

A:Accession: T21947
 A:Reference number: Z19492
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-210 <WIL>
 A:Cross-references: UNIPROT:Q20140; EMBL:Z50045; PI
 A:Experimental source: clone F38B2

A;Gene: CESP:F38B2.4
A;Map position: X
A;Introns: 81/3; 121/3; 149/2; 185/3
C;Superfamily: adenylate kinase

Query Match	100.0%	Score 29;	DB 2;	Length 210;
Best Local Similarity	100.0%;	Pred. No. 2.3e+02;		
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

Qy 1 GGPGS 5
27 GGPGS 31
pb

RESULT 29

T47842

UNIDYLATE KINASE-like protein - Arabidopsis thaliana

N:Alternate names: protein T209.160

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: T47842

R:Nyakatura, G.; Fattmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichsel
submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24475

A:Accession: T47842

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-210 <N/A>

A:Cross-references: UNIPROT:Q9MIC5; EMBL:AL138658

A:Experimental source: cultivar Columbia; BAC clone T209

C:Genetics:

A:Map position: 3

A:Introns: 10/3; 34/1; 73/2; 125/3; 152/3; 188/3

A:Note: T209.160

C:Superfamily: adenylate kinase

Query Match	100.0%;	Score 29;	DB 2;	Length 210;
Best Local Similarity	100.0%;	Pred. No. 2.3e+02;		
Matches 5;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 1 GGPGS 5
db 34 GGPGS 38

RESULT 30

JC7511
fibroblast growth factor-20 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JC7511
R:Ohmachi, S.; Watanabe, Y.; Mikami, T.; Kusu, N.; Ibi, T.; Akaike, A.; Itoh, N.
Biochem. Biophys. Res. Commun. 277, 355-360, 2000
A:Title: FGF-20, a novel neurotrophic factor, preferentially expressed in the substantia
A:Reference number: JC7511; MUID:20490008; PMID:11032730
A:Contents: Brain
A:Accession: JC7511
A:Molecule type: DNA
A:Residues: 1-212 <OHM>

A;Cross-references: UNIPROT:Q9EST9; DDBJ:AB020021
 C;Comment: This factor is secreted, and plays an important role as a neurotrophic factor
 C;Genetics:
 A;Gene: fgf-20
 C;Keywords: brain; growth factor; neurotrophic factor

Query Match 100.0%; Score 29; DB 2; Length 212;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
 |||||
 Db 50 GPGGS 54

RESULT 31
 A81611
 ribosomal protein L3 CP0100 [imported] - Chlamydomophila pneumoniae (strain AR39)
 C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C;Accession: A81611
 R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
 A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A;Reference number: A81500; MUID:20150255; PMID:10684935
 A;Accession: A81611
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-219 <REA>
 A;Cross-references: UNIPROT:Q9Z7Q7; GB:AE002173; GB:AE002161; NID:g7189033; PIDN:AAF3798
 A;Experimental source: strain AR39, HL cells
 C;Genetics:
 A;Gene: CP0100
 C;Superfamily: Escherichia coli ribosomal protein L3

Query Match 100.0%; Score 29; DB 2; Length 219;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
 |||||
 Db 140 GPGGS 144

RESULT 32
 A88102
 protein W09G10.1 [imported] - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C;Accession: A88102
 R;anonymous, The C. elegans Sequencing Consortium.
 A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes.
 A;Reference number: A75000; MUID:99069613; PMID:9851916
 A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 282, 2012-2018, 1998
 A;Accession: A88102
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-222 <STO>
 A;Cross-references: GB:chr_II; PIDN:AB66109.1; PID:g2315665; GSPDB:GN00020; CESP:W09G10.1
 A;Note: Similar to collagen
 C;Genetics:
 A;Gene: W09G10.1
 A;Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 222;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
 |||||

Db 107 GPGGS 111

RESULT 33

A72056

L3 ribosomal protein - Chlamydomophila pneumoniae (strain CWL029)

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000

C;Accession: A72056

R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: A72056

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-225 <ARN>

A;Cross-references: GB:AE001647; GB:AE001363; NID:g4376920; PIDN:AAD18786.1; PID:g4376949

A;Experimental source: strain CWL029

C;Genetics:

A;Gene: rl3

C;Superfamily: Escherichia coli ribosomal protein L3

Query Match 100.0%; Score 29; DB 2; Length 225;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5

|||||

Db 146 GPGGS 150

RESULT 34

D86571

L3 ribosomal protein [imported] - Chlamydomophila pneumoniae (strain J138)

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C;Accession: D86571

R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise,

Nucleic Acids Res. 28, 2311-2314, 2000

A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A;Reference number: A86491; MUID:20330349; PMID:10871362

A;Accession: D86571

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-225 <STO>

A;Cross-references: GB:BA000008; NID:g8979019; PIDN:BAA98954.1; GSPDB:GN00142

A;Experimental source: strain J138

C;Genetics:

A;Gene: rl3

C;Superfamily: Escherichia coli ribosomal protein L3

Query Match 100.0%; Score 29; DB 2; Length 225;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5

|||||

Db 146 GPGGS 150

RESULT 35

C87595

isochorismatase family protein [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: C87595

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay,

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.;

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-239 <ROU>
A;Cross-references: UNIPROT:O80511; EMBL:AC003672; NID:g3341671; PID:g3341692
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
cuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84882
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-239 <STO>
A;Cross-references: GB:AE002093; NID:g3341692; PIDN:AAC27474.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g4720; F16B22.21
A;Map position: 2
A;Introns: 117/3; 152/2; 187/1; 210/3

Query Match 100.0%; Score 29; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 233 GPGS 237

Search completed: August 19, 2005, 16:26:05
Job time : 41 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:15:28 ; Search time 168 Seconds

(without alignments)
15.240 Million cell updates/sec

Title: TUNGAP-CLAIM8

Perfect score: 29

Sequence: 1 99pgs 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	15	Q9UCY7	Q9ucc7 homo sapien
2	29	100.0	16	Q9TCY6	Q9tcy6 oryctolagus
3	29	100.0	31	Q9TKW5	Q9tkw5 mycilus edu
4	29	100.0	49	Q6BQ44	Q6bq44 debaryomyce
5	29	100.0	52	Q6JDN3	Q6jdn3 canis famil
6	29	100.0	58	Q6J371	Q6j371 human immun
7	29	100.0	58	Q76377	Q76377 human immun
8	29	100.0	58	Q76379	Q76379 human immun
9	29	100.0	58	Q76382	Q76382 human immun
10	29	100.0	58	Q76386	Q76386 human immun
11	29	100.0	59	Q76375	Q76375 human immun
12	29	100.0	59	Q76376	Q76376 human immun
13	29	100.0	59	Q76378	Q76378 human immun
14	29	100.0	59	Q76387	Q76387 human immun
15	29	100.0	64	Q23989	Q23989 hordeum vul
16	29	100.0	67	Q9L8X0	Q9l8x0 streptomyce
17	29	100.0	69	Q6LJDJ2	Q6ldj2 homo sapien
18	29	100.0	70	Q9BU27	Q9bu27 homo sapien
19	29	100.0	73	Q6YPC6	Q6ypc6 oryza sativ
20	29	100.0	76	Q6ABQ8	Q6abq8 propionibac
21	29	100.0	81	Q9UKC0	Q9ukc0 homo sapien
22	29	100.0	85	Q9TS70	Q9ts70 bos taurus
23	29	100.0	87	Q77640	Q77640 macaca mula
24	29	100.0	96	Q98UF3	Q98uf3 xenopus lae
25	29	100.0	101	Q95N88	Q95n88 ovis aries
26	29	100.0	102	Q98441	Q98441 paramesicium
27	29	100.0	103	Q9VK53	Q9vks3 drosophila
28	29	100.0	104	Q9SEF3	Q9sef3 hordeum vul
29	29	100.0	107	Q9LUH9	Q9luh9 arabidopsis
30	29	100.0	109	Q9BGZ9	Q9bgz9 macaca fasc
31	29	100.0	110	Q8LT35	Q8lt35 bacterioph

32	29	100.0	110	2	Q8CEP4	Q8cep4 mus musculu
33	29	100.0	112	2	Q17338	Q17338 caenorhabdi
34	29	100.0	112	2	Q20277	Q20277 caenorhabdi
35	29	100.0	114	2	Q9Y8W6	Q9y8w6 aeropyrum p
36	29	100.0	116	2	Q8ITN0	Q8itn0 drosophila
37	29	100.0	120	2	Q9BIU5	Q9biu5 argiope tri
38	29	100.0	121	2	Q7QAR9	Q7qar9 anopheles g
39	29	100.0	126	2	Q8C3U4	Q8c3u4 mus musculu
40	29	100.0	126	2	Q32938	Q32938 mycobacteri
41	29	100.0	127	2	Q67V96	Q67v96 oryza sativ
42	29	100.0	128	2	Q67V96	Q67v96 oryza sativ
43	29	100.0	129	1	ANPB_SHEEP	Q46541 ovis aries
44	29	100.0	131	2	Q69934	Q69934 streptomyce
45	29	100.0	131	2	Q9L089	Q9l089 streptomyce
46	29	100.0	136	2	Q9YG35	Q9yg35 aeropyrum p
47	29	100.0	138	2	Q75AR4	Q75ar4 ashbya goss
48	29	100.0	138	2	Q74993	Q74993 human immun
49	29	100.0	138	2	Q74994	Q74994 human immun
50	29	100.0	139	2	Q9LIZ6	Q9liz6 pinus taeda
51	29	100.0	141	2	Q95KD3	Q95kd3 macaca fasc
52	29	100.0	142	2	Q6Q7K2	Q6q7k2 oryctolagus
53	29	100.0	142	2	Q827K4	Q827k4 streptomyce
54	29	100.0	142	2	Q74967	Q74967 human immun
55	29	100.0	143	1	MK_HUMAN	P21741 homo sapien
56	29	100.0	143	2	Q9N0E6	Q9n0e6 bos taurus
57	29	100.0	143	2	Q74968	Q74968 human immun
58	29	100.0	143	2	Q74996	Q74996 human immun
59	29	100.0	145	2	Q9RKY5	Q9rky5 streptomyce
60	29	100.0	146	2	Q64874	Q64874 arabidopsis
61	29	100.0	148	2	Q7NM01	Q7nm01 gloeobacter
62	29	100.0	148	2	Q84N15	Q84n15 zea mays (m
63	29	100.0	150	2	Q9LGM3	Q9lgm3 oryza sativ
64	29	100.0	152	1	M165_ARATH	Q8c717 arabidopsis
65	29	100.0	152	2	Q8C2F0	Q8c2f0 mus musculu
66	29	100.0	154	2	Q27522	Q27522 methanobact
67	29	100.0	154	2	Q7QF01	Q7qful anopheles g
68	29	100.0	154	2	Q9ZTP6	Q9ztp6 lithospermu
69	29	100.0	154	2	Q6H996	Q6h996 tt virus. h
70	29	100.0	155	2	Q9SEF4	Q9sef4 hordeum vul
71	29	100.0	156	2	Q8LSH6	Q8lsh6 zea mays (m
72	29	100.0	156	2	Q40154	Q40154 lithospermu
73	29	100.0	157	2	Q8MKD4	Q8mkd4 bos taurus
74	29	100.0	158	2	Q7Y083	Q7y083 datisca glo
75	29	100.0	159	2	Q7Y083	Q7y083 datisca glo
76	29	100.0	160	2	Q9MB25	Q9mb25 vigna ungui
77	29	100.0	162	1	ML31_ARATH	Q941r5 arabidopsis
78	29	100.0	162	2	Q8BTI3	Q8bti3 m mus muscu
79	29	100.0	166	2	Q9XHP7	Q9xhp7 ceratopter
80	29	100.0	166	2	Q84QC7	Q84qc7 hordeum vul
81	29	100.0	167	2	Q98M20	Q98m20 rhizobium l
82	29	100.0	169	2	Q77NH9	Q77nh9 rhesus monk
83	29	100.0	169	2	Q9WRM9	Q9wrn9 macaca mula
84	29	100.0	170	2	Q919A1	Q919a1 gallus gall
85	29	100.0	171	2	Q8RXL9	Q8rxl9 arabidopsis
86	29	100.0	171	2	Q8BT81	Q8bt81 mus musculu
87	29	100.0	172	2	Q82N07	Q82n07 streptomyce
88	29	100.0	176	2	Q9BQ61	Q9bq61 homo sapien
89	29	100.0	178	2	Q75F97	Q75f97 ashbya goss
90	29	100.0	179	2	Q8TQ97	Q8tq97 methanosarc
91	29	100.0	181	1	P91207	P91207 caenorhabdi
92	29	100.0	181	1	KAD2_SYNY3	P72641 synechocyst
93	29	100.0	181	2	Q86DY2	Q86dy2 schistosoma
94	29	100.0	182	2	Q744N3	Q744n3 mycobacteri
95	29	100.0	182	2	Q9YC04	Q9yc04 aeropyrum p
96	29	100.0	183	2	Q9VRI3	Q9vri3 drosophila
97	29	100.0	183	2	Q8VPY0	Q8vpy0 serratia ma
98	29	100.0	183	2	Q98KH7	Q98kh7 rhizobium l
99	29	100.0	184	2	Q7PWL7	Q7pwl7 anopheles g
100	29	100.0	184	2	Q711C8	Q711c8 lactobacill

ALIGNMENTS

```

RESULT 1
Q9UCC7 PRELIMINARY; PRT; 15 AA.
AC Q9UC7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Midkine (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=94059921; PubMed=8241100;
RA Novotny W.F., Maffi T., Mehta R.L., Milner P.G.;
RT "Identification of novel heparin-releasable proteins, as well as the
RT cytokines midkine and pleiotrophin, in human postheparin plasma.";
RL Arterioscler. Thromb. 13:1798-1805(1993).
SQ SEQUENCE 15 AA; 1527 MW; C34B6978787474AC CRC64;

Query Match 100.0%; Score 29; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
11 GPGGS 15

RESULT 2
Q9TQY6 PRELIMINARY; PRT; 16 AA.
AC Q9TQY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Annexin I (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=96226532; PubMed=8967522;
RA Mayran N., Traverso V., Maroux S., Massey-Harroche D.;
RT "Cellular and subcellular localizations of annexins I, IV, and VI in
RT lung epithelia.";
RL Am. J. Physiol. 270:L863-L871(1996).
DR HSP; P19619; ILM6.
SQ SEQUENCE 16 AA; 1546 MW; 70AA0A45193C8BB6 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
2 GPGGS 6

RESULT 3
Q9TWK5 PRELIMINARY; PRT; 31 AA.
AC Q9TWK5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Proximal collagen, COL-P (Fragment).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.

```

```

OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=95230211; PubMed=7714453;
RA Qin X., Waite J.H.;
RT "Exotic collagen gradients in the byssus of the mussel Mytilus
RT edulis.";
RL J. Exp. Biol. 198:633-644(1995).
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
FT NON_TER 1 1
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 2648 MW; B1F7708959101A73 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
13 GPGGS 17

RESULT 4
O6BQ44 PRELIMINARY; PRT; 49 AA.
ID O6BQ44
AC O6BQ44;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Debaryomyces hansenii chromosome E of strain CBS767 of Debaryomyces
DE hansenii.
GN ORFNames=DEHA0E09031g;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Ostas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swemene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382137; CAG87912.1; -.
SQ SEQUENCE 49 AA; 5450 MW; 650ECF1C9B63B40A CRC64;

Query Match 100.0%; Score 29; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
3 GPGGS 7

```


RESULT 5

Q6JDN3
ID Q6JDN3 PRELIMINARY; PRT; 52 AA.
AC Q6JDN3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Annexin I (Fragment).
GN Name=ANX1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15233990; DOI=10.1016/j.ygeno.2004.04.001;
RA Housley D.J.E., Ritzert E., Venta P.J.;
RT "Comparative radiation hybrid map of canine chromosome 1 incorporating
RT SNP and indel polymorphisms."
RL Genomics 84:248-264(2004).
DR EMBL; AY514684; AAT44542.1; -.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5536 MW; 1941066A74F911AD CRC64;

Query Match 100.0%; Score 29; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 22 GGPGS 26

RESULT 6

Q76371
ID Q76371 PRELIMINARY; PRT; 58 AA.
AC Q76371;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy."
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26415; AAA53594.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR00777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; AIDS; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 58
SQ SEQUENCE 58 AA; 6312 MW; 0218519B35DD04C CRC64;

Query Match 100.0%; Score 29; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 22 GGPGS 26

Qy 1 GGPGS 5
Db 31 GGPGS 35

RESULT 7

Q76377
ID Q76377 PRELIMINARY; PRT; 58 AA.
AC Q76377;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy."
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26412; AAA53591.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR00777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; AIDS; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 58
SQ SEQUENCE 58 AA; 6384 MW; 002587DC013DD04C CRC64;

Query Match 100.0%; Score 29; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 31 GGPGS 35

RESULT 8

Q76379
ID Q76379 PRELIMINARY; PRT; 58 AA.
AC Q76379;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy."
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26414; AAA53593.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR00777; GP120.

Query Match 100.0%; Score 29; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 31 GGPGS 35

Qy 1 GGPGS 5
Db 31 GGPGS 35

```

DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON TER 1
FT NON TER 58
SQ SEQUENCE 58 AA; 6298 MW; 0748519E650DD04C CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
31 GPGGS 35

RESULT 9
Q76382 PRELIMINARY; PRT; 58 AA.
AC Q76382;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy.";
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26417; AAA53596.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR00777; GP120.
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON TER 1
FT NON TER 58
SQ SEQUENCE 58 AA; 6298 MW; 0748519E650DD04C CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
31 GPGGS 35

RESULT 10
Q76386 PRELIMINARY; PRT; 58 AA.
AC Q76386;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;

```

```

RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy.";
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26416; AAA53595.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR00777; GP120.
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON TER 1
FT NON TER 58
SQ SEQUENCE 58 AA; 6308 MW; 05033C2AF50DD04C CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
31 GPGGS 35

RESULT 11
Q76375 PRELIMINARY; PRT; 59 AA.
AC Q76375;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy.";
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26410; AAA53589.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR00777; GP120.
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON TER 1
FT NON TER 59
SQ SEQUENCE 59 AA; 6542 MW; 145025919E650DD0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
Db |||||
31 GPGGS 35

RESULT 12
Q76376 PRELIMINARY; PRT; 59 AA.
AC Q76376;

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy.";
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26411; AA53592.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 59
SQ SEQUENCE 59 AA; 6570 MW; 145037266B650DD0 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 31 GPGS 35

RESULT 13
Q76378 PRELIMINARY; PRT; 59 AA.
ID O23989 PRELIMINARY; PRT; 59 AA.
AC O23989;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy.";
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26413; AA53592.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 59
SQ SEQUENCE 59 AA; 6584 MW; 14557587D9513DD0 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 31 GPGS 35

RESULT 14
Q76387 PRELIMINARY; PRT; 59 AA.
ID O23989 PRELIMINARY; PRT; 59 AA.
AC O23989;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94201744; PubMed=8151287;
RA Ball J.K., Holmes E.C., Whitwell H., Desselberger U.;
RT "Genomic variation of human immunodeficiency virus type 1 (HIV-1):
RT molecular analyses of HIV-1 in sequential blood samples and various
RT organs obtained at autopsy.";
RL J. Gen. Virol. 75:67-79(1994).
DR EMBL; L26418; AA53597.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 59
SQ SEQUENCE 59 AA; 6486 MW; 2E6218519B350DCA CRC64;

Query Match 100.0%; Score 29; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 31 GPGS 35

RESULT 15
Q23989 PRELIMINARY; PRT; 64 AA.
ID O23989 PRELIMINARY; PRT; 64 AA.
AC O23989;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Expressed sequence tag (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=alobostrians; TISSUE=White leaf;
RA Hees W.R., Golz R.R., Boerner T.;
RT "Analysis of randomly selected cDNAs reveals the expression of
RT stress- and defence related genes in the barley mutant alobostrians.";
RL Plant Sci. 133:191-201(1998).
DR EMBL; AJ000232; CAA03958.1; -.
DR PIR; T05913; T05913.
FT NON_TER 1
FT NON_TER 64
SQ SEQUENCE 64 AA; 6965 MW; 67F33A2A54E6E3DA CRC64;

```

```

Query Match      100.0%; Score 29; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 13 GPGS 17

RESULT 16
Q9L8X0 PRELIMINARY; PRT; 67 AA.
AC Q9L8X0
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyketide synthase module 5 (Fragment).
OS Streptomyces narbonensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=67333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC19790;
RX MEDLINE=20179700; PubMed=10713461; DOI=10.1016/S0378-1119(00)00003-2;
RA Xue Y., Wilson D., Sherman D.H.;
RT "Genetic architecture of the polyketide synthases for methymycin and
RT pikromycin series macrolides.";
RL Gene 245:203-211(2000)
DR EMBL; AF193252; AAF61861.1; -.
FT NON TER
SQ SEQUENCE 67 AA; 6841 MW; B7358A8954E4ED59 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 30 GPGS 34

RESULT 17
Q6LDJ2 PRELIMINARY; PRT; 69 AA.
AC Q6LDJ2
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type III pro-collagen (Fragment).
GN Name=COL3A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dermis;
RX MEDLINE=91161621; PubMed=1672129;
RA Lee B., Vitale E., Superti-Furga A., Steinmann B., Ramirez F.;
RT "G to T transversion at position +5 of a splice donor site causes
RT skipping of the preceding exon in the type III procollagen transcripts
RT of a patient with Ehlers-Danlos syndrome type IV.";
RL J. Biol. Chem. 266:5256-5259(1991).
DR EMBL; M59312; AAA52041.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
KW Collagen.
FT NON TER
SQ SEQUENCE 69 AA; 6219 MW; 936490732D4FDA57 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 7 GPGS 11

RESULT 18
Q9BU27 PRELIMINARY; PRT; 70 AA.
AC Q9BU27
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FAM3A protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Butterfield Y.S.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002934; AAH02934.1; -.
SQ SEQUENCE 70 AA; 7503 MW; 26F4F9C9BC96B7E9 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 28 GPGS 32

RESULT 19
Q6YPC6 PRELIMINARY; PRT; 73 AA.
AC Q6YPC6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSUNOal74H12.9 (Hypothetical protein
DE P0104B02.4).
GN Name=OSUNOal74H12.9; Synonyms=P0104B02.4;
OS Oryza sativa (japonica cultivar-group).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Fujisawa M.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006845; BAD17831.1; -.
DR EMBL; AP006461; BAD10764.1; -.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 7863 MW; 42DAA9C985C26513 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 68 GPGGS 72

RESULT 20
Q6ABQ8
ID Q6ABQ8 PRELIMINARY; PRT; 76 AA.
AC Q6ABQ8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PPA0111;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=KPA171202 / DSM 16379;
RC PubMed=15286373; DOI=10.1126/science.1100330;
RX Brueggemann H., Henne A., Hoster F., Liesegang H., Wierer A.,
RA Strittmatter A., Hujer S., Huerle P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal
of human skin.";
RL Science 305:671-673(2004).
DR EMBL; AE017283; AAT81871.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 76 AA; 7901 MW; ASE12B0595E3CAE7 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 55 GPGGS 59

RESULT 21
Q9UKC0
ID Q9UKC0 PRELIMINARY; PRT; 81 AA.
AC Q9UKC0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE F-box protein Fbx22 (Fragment).
GN Name=FBX22;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20003060; PubMed=10531035; DOI=10.1016/S0960-9822(00)80020-2;
RA Cenciarelli C., Chiur D.S., Guardavaccaro D., Parks W., Vidal M.,
RA Pagano M.;
RT "Identification of a family of human F-box proteins.";
RL Curr. Biol. 9:1177-1179(1999).
DR EMBL; AF174602; AAF04523.1; -.
DR GO; GO:0004840; Fubiquitin conjugating enzyme activity; TAS.
DR GO; GO:0004842; Fubiquitin-protein ligase activity; TAS.
DR GO; GO:0006464; P:protein modification; TAS.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; TAS.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR NON_TER 81
SQ SEQUENCE 81 AA; 8768 MW; B8398FFC30C6CF4B CRC64;

Query Match 100.0%; Score 29; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 8 GPGGS 12

RESULT 22
Q9TS70
ID Q9TS70 PRELIMINARY; PRT; 85 AA.
AC Q9TS70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Midkine, MK (Fragments).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=95163784; PubMed=7859927; DOI=10.1016/0303-7207(94)90171-6;
RA Ohyama Y., Miyamoto K., Minamino N., Matsuo H.;
RT "Isolation and identification of midkine and pleiotrophin in bovine
follicular fluid.";
RL Mol. Cell. Endocrinol. 105:203-208(1994).
DR HSSP; P21741; IMKN.
DR GO; GO:0008083; F:growth factor activity; IEA.
DR InterPro; IPR000762; PTN_MK.
DR PRINTS; PR00269; PTNMDKINE.
DR ProDom; PD005592; PTN_MK; 1.
DR SMART; SM00193; PTN; 1.
DR PROSITE; PS00619; PTN_MK_1; 1.
FT NON_TER 1
FT NON_CONS 47
FT NON_CONS 68
FT NON_TER 85
SQ SEQUENCE 85 AA; 9112 MW; EF44B9784D7BE464 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGGS 5
Db 1 GPGGS 5

RESULT 23
O77640
ID O77640 PRELIMINARY; PRT; 87 AA.

```

```

AC O77640;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dopamine beta-hydroxylase (Fragment).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Ovary.
RC MEDLINE=98393758; PubMed=9724817; DOI=10.1073/pnas.95.18.10990;
RA Mayerhofer A., Smith G.D., Danilchik M., Levine J.E., Wolf D.P.,
RA Dissen G.A., Ojeda S.R.;
RT "Oocytes are a source of catecholamines in the primate ovary: evidence
RT for a cell-cell regulatory loop.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10990-10995(1998).
DR EMBL; AF070919; AAC33898.1; -.
DR GO; GO:000507; F:copper ion binding; IEA.
DR GO; GO:0004500; F:dopamine beta-monooxygenase activity; IEA.
DR GO; GO:0006584; P:catecholamine metabolism; IEA.
DR InterPro; IPR000323; Cu2_monooxygenase.
DR InterPro; IPR000945; DB_monooxygenase.
DR InterPro; IPR008977; PHM_PNGase_F.
DR Pfam; PF01082; Cu2_monooxygenase; I.
DR PRINTS; PR00767; DEMONOXGNASE.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9905 MW; 2F81DBA2DF3CAD70 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
DB 37 GPGGS 41

RESULT 24
Q98UF3 PRELIMINARY; PRT; 96 AA.
ID Q98UF3
AC Q98UF3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Collagen alpha1(III) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RN SEQUENCE FROM N.A.
RP Amano T., Yoshizato K.;
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF170319; AAK11495.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; I.
DR ProDom; PD000007; Clg_helix; I.
DR Collagen.
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 8884 MW; 3D00A45E835F4381 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GPGGS 5
DB 58 GPGGS 62

RESULT 25
Q9SN88 PRELIMINARY; PRT; 101 AA.
ID Q9SN88
AC Q9SN88;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Integrin alpha 5 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
[1]
RN SEQUENCE FROM N.A.
RP Ka H., Johnson G.A., Bazer F.W.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF349459; AAK52480.1; -.
DR HSP; P06756; IL5G.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
KW Integrin.
FT NON_TER 1
FT NON_TER 101
SQ SEQUENCE 101 AA; 11073 MW; 7D6BB57D79946999 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGGS 5
DB 91 GPGGS 95

RESULT 26
Q98441 PRELIMINARY; PRT; 102 AA.
ID Q98441
AC Q98441;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE A389R protein.
GN Name=a389R;
OS Paramecium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96400190; PubMed=8806566; DOI=10.1006/viro.1996.0482;
RX Kutish G.F., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;
RA "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
RT positions 182 to 258.";
RL Virolgy 223:303-317(1996).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20013326; PubMed=10544099; DOI=10.1006/viro.1999.9972;
RX Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Lisec A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homospesmidine
RT synthase.";
RL Virology 263:254-262(1999).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=20478054; PubMed=11021991; DOI=10.1006/viro.2000.0500;
RX Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RA "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT PBCV-1.";

```

```

RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Gurnon J.R.;
RA Submitted (SFP-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; U42580; AAC96757.1; -
DR PIR; T17892; T17892.
SQ SEQUENCE 102 AA; 11452 MW; 543DB74D8DAAB590 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 52 GPGS 56

RESULT 27
Q9VKS3 PRELIMINARY; PRT; 103 AA.
AC
Q9VKS3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG17105-PA.
GN ORFNames=CG17105;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.R., Pacleab J.M.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleab J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J., Wu D.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleab J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003629; AAF52987.1; -
DR FlyBase; FBgn0032280; CG17105.
SQ SEQUENCE 103 AA; 9318 MW; D74CF566E217056D CRC64;

Query Match 100.0%; Score 29; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 69 GPGS 73

RESULT 28
Q9SEF3 PRELIMINARY; PRT; 104 AA.
AC Q9SEF3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Resistance gene analog PIC25-1 (Fragment).
GN Name=PIC25-1;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21337598; PubMed=11444696; DOI=10.1139/gen-44-3-375;
RA Collins N., Park R., Spielmeier W., Ellis J., Pryor A.J.;

```

```

RT "Resistance gene analogs in barley and their relationship to rust
RT resistance genes.",
RL Genome 44:375-381(2001).
DR EMBL; AF146273; AAF21366.1; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON_TER 1
FT NON_TER 104
SQ SEQUENCE 104 AA; 11434 MW; AF18E70529F43A27 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGS 5
DB 27 GPGS 31

RESULT 29
Q9LUH9 PRELIMINARY; PRT; 107 AA.
AC Q9LUH9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MW123
DE (At3g22820).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asanizu E., Tabata S.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022223; BAB01257.1; -.
DR EMBL; BT014813; AAT41796.1; -.
DR EMBL; BT012590; AAT06409.1; -.
SQ SEQUENCE 107 AA; 11505 MW; CSF5E71358745F3A CRC64;

Query Match 100.0%; Score 29; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGS 5
DB 55 GPGS 59

RESULT 30
Q9BGZ9 PRELIMINARY; PRT; 109 AA.
ID Q9BGZ9
AC Q9BGZ9;

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Sugano S.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055257; BAB21881.1; -.
KW Hypothetical protein.
SQ SEQUENCE 109 AA; 11632 MW; 9B9032806827A321 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGS 5
DB 28 GPGS 32

RESULT 31
Q8LT35 PRELIMINARY; PRT; 110 AA.
ID Q8LT35
AC Q8LT35;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Restriction endonuclease.
OS Bacteriophage MB78 (Salmonella typhimurium bacteriophage MB78).
OC Viruses.
OX NCBI_TaxID=52971;
RN [1]
RP SEQUENCE FROM N.A.
RA Chaturvedi D., Chakravorty M.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277754; CAC81910.1; -.
DR HSPB; P12528; 1TYV.
DR GO; GO:0004519; P:endonuclease activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
KW Endonuclease.
SQ SEQUENCE 110 AA; 11507 MW; 5BD83A908D7DA6A1 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGS 5
DB 87 GPGS 91

RESULT 32
Q8CEP4 PRELIMINARY; PRT; 110 AA.
ID Q8CEP4
AC Q8CEP4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched
DE library, clone:543041J08 product:hypothetical protein, full insert
DE sequence. (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```



```
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK017296; BAC25509.1; -
KW Hypothetical protein.
FT NON_TER
RP SEQUENCE 110 AA; 11536 MW; 0F5398A6AE90177 CRC64;
SQ SEQUENCE 110 AA; 11536 MW; 0F5398A6AE90177 CRC64;
```

```
Query Match 100.0%; Score 29; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GPGFS 5
Db 63 GPGFS 67
```

```
RESULT 33
Q17338 PRELIMINARY; PRT; 112 AA.
AC Q17338;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to enteropeptidase (Fragment).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95262900; PubMed=7744249;
RA Varkey J.P., Muhirad P.J., Minniti A.N., Do B., Ward S.;
RT "The Caenorhabditis elegans spe-26 gene is necessary to form
RT spermatids and encodes a protein similar to the actin-associated
RT proteins kelch and scruin.";
RL Genes Dev. 9:1074-1086(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=93078873; PubMed=1448167; DOI=10.1038/360456a0;
RA Van Voorhies W.A.;
RT "Production of sperm reduces nematode lifespan.";
RL Nature 360:456-458(1992).
DR EMBL; U20142; AAB60234.1; -
DR InterPro; IPR000859; CUB.
DR Pfam; PF00431; CUB; 1.
DR SMART; SM00042; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
FT NON_TER
FT NON_TER 112
SQ SEQUENCE 112 AA; 12166 MW; E95B496A3C296697 CRC64;
Query Match 100.0%; Score 29; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPGFS 5
Db 49 GPGFS 53
```

```
RESULT 34
Q20277 PRELIMINARY; PRT; 112 AA.
AC Q20277;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein F41E7.5.
GN ORFNames=F41E7.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lennard N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68106; CAA92128.1; -
```

```

DR PIR; T22078; T22078.
DR WormBase; WBGene00009621; F41E7.5.
DR WormPep; F41E7.5; CE03305.
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 10619 MW; 7505B58C593A47C6 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 51 GPGS 55

RESULT 35
Q9Y8W6 PRELIMINARY; PRT; 114 AA.
AC Q9Y8W6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE2518.
GN OrderedLocusNames=APE2518;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococccaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KJ;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000064; BAA81534.1; -.
DR PIR; F72484; F72484.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 114 AA; 12481 MW; 6D5E818EF877785C CRC64;

Query Match      100.0%; Score 29; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 39 GPGS 43

RESULT 36
Q8ITN0 PRELIMINARY; PRT; 116 AA.
AC Q8ITN0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Centromeric histone (Fragment).
GN Name=cid; Synonyms=Cid;
OS Drosophila rajasekari.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=192847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KJ;
RX Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RA Malik H.S., Vermaak D., Henikoff S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR PIR; T22078; T22078.
DR WormBase; WBGene00009621; F41E7.5.
DR WormPep; F41E7.5; CE03305.
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 10619 MW; 7505B58C593A47C6 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 51 GPGS 55

RESULT 35
Q9Y8W6 PRELIMINARY; PRT; 114 AA.
AC Q9Y8W6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE2518.
GN OrderedLocusNames=APE2518;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococccaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KJ;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000064; BAA81534.1; -.
DR PIR; F72484; F72484.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 114 AA; 12481 MW; 6D5E818EF877785C CRC64;

Query Match      100.0%; Score 29; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 39 GPGS 43

RESULT 36
Q8ITN0 PRELIMINARY; PRT; 116 AA.
AC Q8ITN0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Centromeric histone (Fragment).
GN Name=cid; Synonyms=Cid;
OS Drosophila rajasekari.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=192847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KJ;
RX Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RA Malik H.S., Vermaak D., Henikoff S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF435355; AAN63610.1; -.
DR FlyBase; FBgn0064044; Draj\cid.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 12861 MW; DD0CDAD0DD8AA979 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 28 GPGS 32

RESULT 37
Q9BIU5 PRELIMINARY; PRT; 120 AA.
AC Q9BIU5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major ampullate spidroin 2-like protein (Fragment).
OS Argiope trifasciata (Banded garden spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneidae; Argiope.
OX NCBI_TaxID=156845;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21179804; PubMed=11283372; DOI=10.1126/science.1057561;
RA Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;
RT "Extreme diversity, conservation, and convergence of spider silk
RT fibroin sequences.";
RL Science 291:2603-2605(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Gatesy J.E., Hayashi C.Y.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350268; AAK30597.1; -.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 9726 MW; D4DE730C0BA02983 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 62 GPGS 66

RESULT 38
Q7QAR9 PRELIMINARY; PRT; 121 AA.
AC Q7QAR9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AgCP7860.
GN Name=agCG49510; ORFNames=ENSANGG00000008543;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RX Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC      preliminary data.
DR      EMBL; AAB01008888; EAA08868.1; -.
DR      HSP; P04058; 1EVE.
DR      InterPro; IPR002018; CarboxylaseB.
DR      Pfam; PF00135; Coesterase; 1.
SQ      SEQUENCE 121 AA; 12642 MW;  FA4EA9EB8745A57C0 CRC64;

Query Match      100.0%; Score 29; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      34 GGPGS 38

RESULT 39
Q8C3U4      PRELIMINARY; PRT; 126 AA.
ID      Q8C3U4;
AC      Q8C3U4;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched
DE      library, clone:D430006M22 product:hypothetical protein, full insert
DE      sequence. (Fragment).
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Lung;
RX      MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA      Carninci P., Hayashizaki Y.;
RT      "High-efficiency full-length cDNA cloning.";
RL      Mech. Enzymol. 303:19-44(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Lung;
RX      MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA      RIKEN FANTOM Consortium;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Lung;
RA      The FANTOM Consortium;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Lung;
RX      MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA      Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA      Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT      "Normalization and subtraction of cap-trapper-selected cDNAs to
RT      prepare full-length cDNA libraries for rapid discovery of new genes.";
RL      Genome Res. 10:1617-1630(2000).
RN      [5]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Lung;
RX      MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA      Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA      Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA      Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA      Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA      Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA      Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA      Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT      "RIKEN integrated sequence analysis (RISA) system-384-format

```

Search completed: August 19, 2005, 16:25:22

Job time : 173 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:16:58 ; Search time 42 Seconds

(without alignments)
8.887 Million cell updates/sec

Title: TUNGAP-CLAIM8

Perfect score: 29

Sequence: 1 9999s 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA.*
1: /cgm2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgm2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgm2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgm2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgm2_6/ptodata/1/iaa/PTCUS COMB.pep.*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	13	3	US-08-526-136-15
2	29	100.0	35	3	US-09-082-279B-636
3	29	100.0	35	3	US-09-315-304B-636
4	29	100.0	35	4	US-09-834-784-636
5	29	100.0	35	4	US-09-515-965A-636
6	29	100.0	35	4	US-09-350-641C-636
7	29	100.0	35	4	US-08-176-500-130
8	29	100.0	38	1	US-08-471-052A-130
9	29	100.0	38	1	US-08-189-311-130
10	29	100.0	38	2	US-08-471-933-130
11	29	100.0	38	2	US-08-471-800-130
12	29	100.0	38	2	US-08-471-068-130
13	29	100.0	51	1	US-08-253-155A-30
14	29	100.0	51	3	US-08-963-851-13
15	29	100.0	58	4	US-09-513-999C-5823
16	29	100.0	103	4	US-09-270-767-3929
17	29	100.0	111	4	US-09-270-767-3929
18	29	100.0	111	4	US-09-270-767-54146
19	29	100.0	113	3	US-09-220-528-3
20	29	100.0	113	4	US-09-347-613C-12
21	29	100.0	113	4	US-09-662-183A-12
22	29	100.0	116	3	US-09-220-528-4
23	29	100.0	116	4	US-09-347-613C-11
24	29	100.0	116	4	US-09-662-183A-11
25	29	100.0	120	4	US-09-252-991A-22272
26	29	100.0	121	3	US-09-013-084-3
27	29	100.0	133	4	US-09-270-767-34410

28	29	100.0	133	4	US-09-270-767-49627	Sequence 49627, A
29	29	100.0	139	4	US-09-513-999C-5003	Sequence 5003, Ap
30	29	100.0	140	3	US-09-220-528-5	Sequence 5, Appli
31	29	100.0	140	4	US-09-347-613C-10	Sequence 10, Appl
32	29	100.0	140	4	US-09-662-183A-10	Sequence 10, Appl
33	29	100.0	143	3	US-09-012-084-2	Sequence 2, Appli
34	29	100.0	143	4	US-09-380-882-6	Sequence 2, Appli
35	29	100.0	144	1	US-07-874-848B-2	Sequence 2, Appli
36	29	100.0	150	4	US-09-489-039A-13888	Sequence 13888, A
37	29	100.0	159	3	US-09-220-528-12	Sequence 12, Appl
38	29	100.0	159	3	US-09-220-528-89	Sequence 89, Appl
39	29	100.0	161	4	US-09-821-687-2	Sequence 2, Appli
40	29	100.0	166	4	US-09-252-991A-20662	Sequence 20662, A
41	29	100.0	170	4	US-09-252-991A-21784	Sequence 21784, A
42	29	100.0	170	4	US-09-489-039A-13878	Sequence 13878, A
43	29	100.0	181	3	US-09-220-528-40	Sequence 40, Appl
44	29	100.0	186	4	US-09-370-767-60337	Sequence 60337, A
45	29	100.0	186	4	US-09-248-796A-18112	Sequence 18112, A
46	29	100.0	189	4	US-09-543-681A-7932	Sequence 7932, Ap
47	29	100.0	194	2	US-08-879-561-11	Sequence 11, Appl
48	29	100.0	194	2	US-08-879-561-12	Sequence 12, Appl
49	29	100.0	197	2	US-08-879-561-5	Sequence 5, Appli
50	29	100.0	197	4	US-09-252-991A-22880	Sequence 22880, A
51	29	100.0	198	4	US-09-489-039A-10492	Sequence 10492, A
52	29	100.0	198	4	US-09-902-540-12139	Sequence 12139, A
53	29	100.0	202	4	US-09-252-991A-26333	Sequence 26333, A
54	29	100.0	202	4	US-09-949-016-11709	Sequence 11709, A
55	29	100.0	212	4	US-09-692-945-2	Sequence 2, Appli
56	29	100.0	219	4	US-09-893-737-130	Sequence 130, App
57	29	100.0	220	3	US-09-220-528-26	Sequence 26, Appl
58	29	100.0	220	4	US-09-347-613C-9	Sequence 9, Appli
59	29	100.0	220	4	US-09-347-613C-35	Sequence 35, Appl
60	29	100.0	220	4	US-09-662-183A-9	Sequence 9, Appli
61	29	100.0	220	4	US-09-662-183A-35	Sequence 35, Appl
62	29	100.0	235	4	US-09-398-452A-684	Sequence 684, App
63	29	100.0	236	4	US-09-640-211A-1063	Sequence 1063, Ap
64	29	100.0	237	3	US-09-220-528-32	Sequence 32, Appl
65	29	100.0	239	4	US-09-489-039A-7567	Sequence 7567, Ap
66	29	100.0	241	4	US-09-370-767-43033	Sequence 43033, A
67	29	100.0	244	4	US-09-252-991A-31668	Sequence 31668, A
68	29	100.0	249	4	US-09-252-991A-21188	Sequence 21188, A
69	29	100.0	249	4	US-09-949-016-9480	Sequence 9480, Ap
70	29	100.0	260	4	US-09-270-767-43301	Sequence 43301, A
71	29	100.0	269	1	US-08-452-531-4	Sequence 4, Appli
72	29	100.0	269	2	US-08-460-555-4	Sequence 4, Appli
73	29	100.0	269	2	US-08-460-555-4	Sequence 4, Appli
74	29	100.0	269	3	US-08-460-555-4	Sequence 4, Appli
75	29	100.0	269	4	US-09-270-767-43140	Sequence 43140, A
76	29	100.0	275	4	US-09-489-039A-8298	Sequence 8298, Ap
77	29	100.0	278	4	US-09-438-185A-649	Sequence 649, App
78	29	100.0	281	4	US-09-252-991A-25476	Sequence 25476, A
79	29	100.0	289	4	US-09-949-016-6282	Sequence 6282, Ap
80	29	100.0	294	3	US-09-131-028A-9	Sequence 9, Appli
81	29	100.0	294	3	US-09-131-028A-11	Sequence 11, Appl
82	29	100.0	306	4	US-09-949-016-7825	Sequence 7825, Ap
83	29	100.0	310	4	US-09-800-729-190	Sequence 190, App
84	29	100.0	317	4	US-09-738-946-2	Sequence 2, Appli
85	29	100.0	319	4	US-09-902-540-14554	Sequence 14554, A
86	29	100.0	328	4	US-09-900-527-2	Sequence 2, Appli
87	29	100.0	346	4	US-08-948-276-6	Sequence 39, Appl
88	29	100.0	346	4	US-09-919-172-39	Sequence 11617, A
89	29	100.0	348	4	US-09-949-016-11617	Sequence 26571, A
90	29	100.0	356	4	US-09-252-991A-26571	Sequence 26571, A
91	29	100.0	357	4	US-09-252-991A-30141	Sequence 30141, A
92	29	100.0	366	4	US-09-252-991A-23504	Sequence 23504, A
93	29	100.0	370	4	US-09-949-016-11534	Sequence 11534, A
94	29	100.0	385	4	US-09-311-021-78	Sequence 78, Appl
95	29	100.0	392	4	US-08-311-731A-210	Sequence 210, App
96	29	100.0	403	4	US-09-949-016-7412	Sequence 7412, Ap
97	29	100.0	414	1	US-08-255-471-9	Sequence 9, Appli
98	29	100.0	415	4	US-09-252-991A-33056	Sequence 33056, A
99	29	100.0	417	4	US-09-252-991A-24932	Sequence 24932, A
100	29	100.0	420	4	US-09-328-352-7296	Sequence 7296, Ap

ALIGNMENTS

```
RESULT 1
US-08-526-136-15
; Sequence 15, Application US/08526136
; Patent No. 6107089
; GENERAL INFORMATION:
; APPLICANT: Towle, Christine A. et al.
; TITLE OF INVENTION: ANNEXIN XI
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,136
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,036
; FILING DATE:
; APPLICATION NUMBER: 07/837,775
; FILING DATE: February 13, 1992
; APPLICATION NUMBER: 07/764,465
; FILING DATE: September 23, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/099001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
;
US-08-526-136-15
Query Match 100.0%; Score 29; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 5 GGPGS 9

RESULT 2
US-09-082-279B-636
; Sequence 636, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohmed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; PHARMACOKINETIC PROPERTIES
```

```
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-636
Query Match 100.0%; Score 29; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 14 GGPGS 18

RESULT 3
US-09-315-304B-636
; Sequence 636, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; PROPERTIES
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-636
Query Match 100.0%; Score 29; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 14 GGPGS 18

RESULT 4
US-09-834-784-636
; Sequence 636, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohmed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/834,784
; CURRENT FILING DATE: 2001-04-13
```

; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-834-784-636

Query Match 100.0%; Score 29; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 14 GGPGS 18

RESULT 5

US-09-515-965A-636
; Sequence 636, Application US/09515965A
; Patent No. 6623741
; GENERAL INFORMATION:
; APPLICANT: Antczak, J.
; APPLICANT: Delmedico, M.
; APPLICANT: Erickson, J.
; APPLICANT: Lambert, D.
; APPLICANT: Sista, P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE FUSION-
; FILE REFERENCE: 7872-073
; CURRENT APPLICATION NUMBER: US/09/515.965A
; CURRENT FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1994
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-515-965A-636

Query Match 100.0%; Score 29; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 14 GGPGS 18

RESULT 6

US-09-350-641C-636
; Sequence 636, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C

; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-350-641C-636

Query Match 100.0%; Score 29; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 14 GGPGS 18

RESULT 7

US-09-350-841A-636
; Sequence 636, Application US/09350841A
; Patent No. 6750008
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 636
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-350-841A-636

Query Match 100.0%; Score 29; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 14 GGPGS 18

RESULT 8

US-08-176-500-130
; Sequence 130, Application US/08176500
; Patent No. 5498538
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,416
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-143
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-176-500-130

Query Match      100.0%; Score 29; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GPGGS 5
DB      17 GPGGS 21

RESULT 9
US-08-471-052A-130
; Sequence 130, Application US/08471052A
; Patent No. 5625033
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,052A
; FILING DATE: 06-JUNE-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-179
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Query Match      100.0%; Score 29; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GPGGS 5
DB      17 GPGGS 21

RESULT 10
US-08-189-331-130
; Sequence 130, Application US/08189331
; Patent No. 5747334
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 186
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/189,331
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-155
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-189-331-130

Query Match      100.0%; Score 29; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GPGGS 5
DB      17 GPGGS 21

RESULT 11
US-08-471-939-130
; Sequence 130, Application US/08471939
; Patent No. 5844076
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 141
```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,939
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,416
FILING DATE: 01-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-143
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 130:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-471-939-130

Query Match 100.0%; Score 29; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 17 GGPGS 21

RESULT 13
US-08-471-068-130
Sequence 130, Application US/08471068
Patent No. 5948635
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,068
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/189,331
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 130:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-471-068-130

Query Match 100.0%; Score 29; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 17 GGPGS 21

RESULT 12
US-08-471-800-130
Sequence 130, Application US/08471800
Patent No. 5852167
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,800
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/013,416
FILING DATE:

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGPGS 5
Db 17 GGPGS 21

RESULT 14
US-08-253-155A-30
; Sequence 30, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenő
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-253-155A-30

Query Match 100.0%; Score 29; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
Db 13 GGPGS 17

RESULT 15
US-08-963-851-13
; Sequence 13, Application US/08963851
; Patent No. 6300116
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/08/963,851
; CURRENT FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 13
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-08-963-851-13

Query Match 100.0%; Score 29; DB 3; Length 58;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
Db 27 GGPGS 31

RESULT 16
US-09-513-999C-5823
; Sequence 5823, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5823
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 89
; OTHER INFORMATION: Xaa=His or Leu or Pro or Gln or Arg
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 94
; OTHER INFORMATION: Xaa=Asp or Glu
US-09-513-999C-5823

Query Match 100.0%; Score 29; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGPGS 5
Db 35 GGPGS 39

RESULT 17
US-09-270-767-38929
; Sequence 38929, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38929
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid

```
US-09-270-767-38929
Query Match      100.0%; Score 29; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      84 GGPGS 88

RESULT 18
US-09-270-767-54146
; Sequence 54146, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54146
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54146

Query Match      100.0%; Score 29; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      84 GGPGS 88

RESULT 19
US-09-220-528-3
; Sequence 3, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-220-528-3

Query Match      100.0%; Score 29; DB 3; Length 113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      2 GGPGS 6

RESULT 20
US-09-347-613C-12
; Sequence 12, Application US/09347613C
; Patent No. 6593133
; GENERAL INFORMATION:
; APPLICANT: Johansen, Teit E.
; APPLICANT: Blom, Nikolaj
; APPLICANT: Hansen, Claus
; TITLE OF INVENTION: No. 6593133el Neurotrophic Factors
; FILE REFERENCE: NeuroSearch 19313-001
; CURRENT APPLICATION NUMBER: US/09/347,613C
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: DANISH 1998 00904
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: USSN 60/092,229
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: DANISH 1998 01048
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: USSN 60/097,774
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: DANISH 1998 01260
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,908
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: DANISH 1998 01265
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: (95)
; OTHER INFORMATION: glycosylated asparagine
US-09-347-613C-12

Query Match      100.0%; Score 29; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGPGS 5
Db      2 GGPGS 6

RESULT 21
US-09-662-183A-12
; Sequence 12, Application US/09662183A
; Patent No. 6734284
; GENERAL INFORMATION:
; APPLICANT: Johansen, Teit E.
; APPLICANT: Blom, Nikolaj
; APPLICANT: Hansen, Claus
; TITLE OF INVENTION: No. 6734284el Neurotrophic Factors
; FILE REFERENCE: 19313-001 DIV
; CURRENT APPLICATION NUMBER: US/09/662,183A
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: DANISH 1998 00904
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: USSN 60/092,229
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: DANISH 1998 01048
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: USSN 60/097,774
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: DANISH 1998 01260
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,908
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: DANISH 1998 01265
```

; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 09/347,613
; PRIOR FILING DATE: 2000-07-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: (95)
; OTHER INFORMATION: glycosylated asparagine
US-09-662-183A-12

Query Match 100.0%; Score 29; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 2 GGPGS 6

RESULT 22
US-09-220-528-4
; Sequence 4, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-220-528-4

Query Match 100.0%; Score 29; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 5 GGPGS 9

RESULT 23
US-09-347-613C-11
; Sequence 11, Application US/09347613C
; Patent No. 6593133
; GENERAL INFORMATION:
; APPLICANT: Johansen, Teit E.
; APPLICANT: Blom, Nikolaj
; APPLICANT: Hansen, Claus
; TITLE OF INVENTION: No. 6593133el Neurotrophic Factors
; FILE REFERENCE: NeuroSearch 19313-001
; CURRENT APPLICATION NUMBER: US/09/347,613C
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: DANISH 1998 00904
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: USSN 60/092,229

; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: DANISH 1998 01048
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: USSN 60/097,774
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: DANISH 1998 01260
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,908
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: DANISH 1998 01265
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: (98)
; OTHER INFORMATION: glycosylated asparagine
US-09-347-613C-11

Query Match 100.0%; Score 29; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 5 GGPGS 9

RESULT 24
US-09-662-183A-11
; Sequence 11, Application US/09662183A
; Patent No. 6734284
; GENERAL INFORMATION:
; APPLICANT: Johansen, Teit E.
; APPLICANT: Blom, Nikolaj
; APPLICANT: Hansen, Claus
; TITLE OF INVENTION: No. 6734284el Neurotrophic Factors
; FILE REFERENCE: 19313-001 DIV
; CURRENT APPLICATION NUMBER: US/09/662,183A
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: DANISH 1998 00904
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: USSN 60/092,229
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: DANISH 1998 01048
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: USSN 60/097,774
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: DANISH 1998 01260
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,908
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: DANISH 1998 01265
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 09/347,613
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: (98)
; OTHER INFORMATION: glycosylated asparagine
US-09-662-183A-11

Query Match 100.0%; Score 29; DB 4; Length 116;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 5 GGPGS 9

RESULT 25
US-09-252-991A-22272
; Sequence 22272, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22272
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22272

Query Match 100.0%; Score 29; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 107 GGPGS 111

RESULT 26
US-09-012-084-3
; Sequence 3, Application US/09012084
; Patent No. 6083907
; GENERAL INFORMATION:
; APPLICANT: UCHIDA, Masayuki
; APPLICANT: IKEMATSU, Shinya
; APPLICANT: YOKOYAMA, Minehiko
; APPLICANT: YAMASHITA, Akio
; APPLICANT: KUMAI, Hideshi
; APPLICANT: ODA, Munehiro
; APPLICANT: KATO, Naoki
; APPLICANT: SAKUMA, Sadatoshi
; APPLICANT: MURAMATSU, Takashi
; TITLE OF INVENTION: ANTI-ULCER COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,084
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Face, Doran R.

Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 5 GGPGS 9

RESULT 27
US-09-270-767-34410
; Sequence 34410, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34410
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34410

Query Match 100.0%; Score 29; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 45 GGPGS 49

RESULT 28
US-09-270-767-49627
; Sequence 49627, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49627
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49627

Query Match 100.0%; Score 29; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 45 GPGS 49

RESULT 29
US-09-513-999C-5003
; Sequence 5003, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5003:
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5003

Query Match 100.0%; Score 29; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 30 GPGS 34

RESULT 30
US-09-520-528-5
; Sequence 5, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Balch, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-520-528-5

Query Match 100.0%; Score 29; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 45 GPGS 49

Db 29 GPGS 33
RESULT 31
US-09-347-613C-10
; Sequence 10, Application US/09347613C
; Patent No. 6593133
; GENERAL INFORMATION:
; APPLICANT: Johansen, Teit E.
; APPLICANT: Blom, Nikolaj
; APPLICANT: Hansen, Claus
; TITLE OF INVENTION: No. 6593133el Neurotrophic Factors
; FILE REFERENCE: NeuroSearch 19133-001
; CURRENT APPLICATION NUMBER: US/09/347,613C
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: DANISH 1998 00904
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: USSN 60/092,229
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: DANISH 1998 01048
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: USSN 60/097,774
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: DANISH 1998 01260
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,908
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: DANISH 1998 01265
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: (122)
; OTHER INFORMATION: glycosylated asparagine
US-09-347-613C-10

Query Match 100.0%; Score 29; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGS 5
Db 29 GPGS 33

RESULT 32
US-09-662-183A-10
; Sequence 10, Application US/09662183A
; Patent No. 6734284
; GENERAL INFORMATION:
; APPLICANT: Johansen, Teit E.
; APPLICANT: Blom, Nikolaj
; APPLICANT: Hansen, Claus
; TITLE OF INVENTION: No. 6734284el Neurotrophic Factors
; FILE REFERENCE: 19313-001 DIV
; CURRENT APPLICATION NUMBER: US/09/662,183A
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: DANISH 1998 00904
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: USSN 60/092,229
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: DANISH 1998 01048
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: USSN 60/097,774
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: DANISH 1998 01260
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,908

;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: DANISH 1998 01265
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 09/347,613
;; PRIOR FILING DATE: 2000-07-02
;; NUMBER OF SEQ ID NOS: 43
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 10
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CARBOHYD
;; LOCATION: (122)
;; OTHER INFORMATION: glycosylated asparagine
US-09-662-183A-10

Query Match 100.0%; Score 29; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 29 GGPGS 33

RESULT 33
US-09-012-084-2
; Sequence 2, Application US/09012084
; Patent No. 6083907
; GENERAL INFORMATION:
; APPLICANT: UCHIDA, Masayuki
; APPLICANT: IKEMATSU, Shinya
; APPLICANT: YOKOYAMA, Minehiko
; APPLICANT: YAMASHITA, Akio
; APPLICANT: KUMAI, Hideshi
; APPLICANT: ODA, Munehiro
; APPLICANT: KATO, Naoki
; APPLICANT: SAKUMA, Sadatoshi
; APPLICANT: MURAMATSU, Takashi
; TITLE OF INVENTION: ANTI-ULCER COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,084
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Doran R.
; REGISTRATION NUMBER: 38,261
; REFERENCE/DOCKET NUMBER: SPO-102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 143 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:

;; ORGANISM: Homo sapiens
US-09-012-084-2

Query Match 100.0%; Score 29; DB 3; Length 143;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 31 GGPGS 35

RESULT 34
US-09-380-882-6
; Sequence 6, Application US/09380882
; Patent No. 6572851
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Takashi
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Oda, Munehiro
; APPLICANT: Ikematsu, Shinya
; APPLICANT: Sakuma, Sadatoshi
; TITLE OF INVENTION: Preventive and Therapeutic Compositions for Drug-Induced Nephropathy
; FILE REFERENCE: SPO-105
; CURRENT APPLICATION NUMBER: US/09/380,882
; CURRENT FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-380-882-6

Query Match 100.0%; Score 29; DB 4; Length 143;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 31 GGPGS 35

RESULT 35
US-07-874-848B-2
; Sequence 2, Application US/07874848B
; Patent No. 5461029
; GENERAL INFORMATION:
; APPLICANT: Joseph Mark Backer and
; APPLICANT: Michael Robert Ostrander
; TITLE OF INVENTION: Method of treating and
; TITLE OF INVENTION: Preventing Viral Infections
; TITLE OF INVENTION: Using HENF and MK Protein
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Antoinette F. Konski
; ADDRESSEE: American Cyanamid Company
; STREET: 1937 West Main Street
; STREET: P.O. Box 60
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06904-0060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII from IBM Displaywrite 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/874,848B
; FILING DATE: 19920424
; CLASSIFICATION: 514

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Koneki, Antoinette F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 31850-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203 321 2455
; TELEFAX: 203 321 2971
; TELEX: 710 474 4059
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 Amino Acids
; TYPE: AMINO ACID
; TOPOLOGY: Linear
; MOLECULE TYPE: Protein
; US-07-874-848B-2

Query Match 100.0%; Score 29; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 31 GGPGS 35

RESULT 36
US-09-489-039A-13888
; Sequence 13888, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13888
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-13888

Query Match 100.0%; Score 29; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 6 GGPGS 10

RESULT 37
US-09-220-528-12
; Sequence 12, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
```

```
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-220-528-12

Query Match 100.0%; Score 29; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 48 GGPGS 52

RESULT 38
US-09-220-528-89
; Sequence 89, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-220-528-89

Query Match 100.0%; Score 29; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
Db 48 GGPGS 52

RESULT 39
US-09-821-687-2
; Sequence 2, Application US/09821687
; Patent No. 6794500
; GENERAL INFORMATION:
; APPLICANT: MIKOSHIBA, KATSUHIKO
; APPLICANT: MIZUTANI, AKIHIRO
; TITLE OF INVENTION: RNA-BINDING PROTEIN
; FILE REFERENCE: 081356/0162
; CURRENT APPLICATION NUMBER: US/09/821,687
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-299812
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-821-687-2
```


Query Match 100.0%; Score 29; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|
|
|
|
|
Db 127 GGPGS 131

RESULT 40
US-09-252-991A-20662
; Sequence 20662, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20662
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20662

Query Match 100.0%; Score 29; DB 4; Length 166;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGPGS 5
|
|
|
|
|
Db 146 GGPGS 150

Search completed: August 19, 2005, 16:26:52
Job time : 43 secs

This Page Blank (uspto)

99 112 100.0 262 7 ADH44200 Adh44200 sFv antib
100 112 100.0 262 7 ADH44198 Adh44198 sFv antib

ALIGNMENTS

RESULT 1
AAR87737
ID AAR87737 standard; peptide; 20 AA.

AC AAR87737;

DT 28-MAY-1996 (first entry)

DE DE Spacer #5 for a circularly permuted ligand.

XX Linker; interleukin-4; interleukin-2; IL-4; IL-2; G-CSF; GM-CSF;
KW granulocyte-colony stimulating factor; circularly permuted ligand;
KW cancer; therapy; granulocyte macrophage-colony stimulating factor.

XX Synthetic.

XX WO9527732-A2.

XX 19-OCT-1995.

XX 06-APR-1995; 95WO-US004468.

XX 08-APR-1994; 94US-00225224.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Pastan I, Kreitman RJ;

XX WPI; 1995-366354/47.

XX Circularly permuted ligands and chimeric mols. are modified
PT interleukin(s) or colony-stimulating factors - with increased binding
PT specificity and affinity for inhibiting tumours.

XX Example 3; Page 42; 97pp; English.

XX This sequence represents a spacer for a modified ligand. The modified
CC ligand is constructed from two copies of a sequence joined to a linker.
CC The sequences used in the ligand include interleukin-4 (IL-4), IL-2,
CC granulocyte-colony stimulating factor (G-CSF) and granulocyte macrophage-
CC colony stimulating factor (GM-CSF). The linker sequences are represented
CC by AAR87726 and AAR87730. These modified ligands are known as circularly
CC permuted ligands. This sequence is used to join the circularly permuted
CC ligand to a protein. The protein used is preferably a cytotoxin, or an
CC antibody. The circularly permuted ligands have greater specificity and
CC binding affinity than the original ligand. They can be used in chimeric
CC molecules to inhibit the growth of tumour cells in vivo, where the two
CC copies of the protein sequence are joined to cytotoxins

XX Sequence 20 AA;

Query Match 100.0%; Score 112; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 2
AAW18554
ID AAW18554 standard; protein; 20 AA.

XX AAW18554;

XX

DT 27-MAR-1998 (first entry)
XX Linker sequence 1 for use in a novel fusion protein.
XX Kunitz type active site inhibitor domain; fusion protein; Factor VIIa;
KW FVIIa; linker; tissue factor; TF; treatment; procoagulant activity;
KW thrombosis.
XX Synthetic.
XX WO9720939-A1.
XX 12-JUN-1997.
XX 22-NOV-1996; 96WO-US018756.
XX 01-DEC-1995; 95US-00566459.
PR 04-DEC-1995; 95US-00566800.
XX (GETH) GENENTECH INC.
PA Kelley RF, Lazarus RA, Lee GF;
XX WPI; 1997-319785/29.
XX Factor VIIa active site inhibitor domain and tissue factor fusion protein
PT - used for inhibiting human tissue factor-FVIIa procoagulant activity in
PT mammals, to treat e.g. thrombosis and restenosis.

XX Claim 6; Page 43; 58pp; English.

XX This linker sequence is used to construct a novel fusion protein. The
CC novel fusion protein comprises a FVIIa active site inhibitor domain (the
CC kunitz domain), a linker domain, and a tissue factor (TF) domain. The
CC novel fusion protein can be used for inhibiting human tissue factor-FVIIa
CC (TF-FVIIa) procoagulant activity in a mammal. It can be used to treat TF-
CC FVIIa related disease-like chronic thromboembolic diseases or disorders
CC associated with fibrin formation including vascular disorders such as
CC deep venous thrombosis, arterial thrombosis, stroke, tumour metastasis,
CC thrombolysis, arteriosclerosis and restenosis following angioplasty. The
CC novel fusion protein can also be used to treat acute and chronic
CC disorders such as inflammation, septic shock, adult respiratory distress
CC syndrome, septicemia, hypotension, disseminated intravascular
CC coagulopathy and other diseases. Fusion proteins in which a FVIIa active
CC site inhibitor and a tissue factor are covalently linked are more potent
CC inhibitors of FVIIa than the individual components alone or a non-
CC covalent mixture of the two

XX Sequence 20 AA;

Query Match 100.0%; Score 112; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 3
AAV43415
ID AAV43415 standard; peptide; 20 AA.

XX AAV43415;

XX 20-DEC-1999 (first entry)

DE Peptide SEQ ID NO:14.

XX Angiogenic homology region; AHR; thrombospondin 1; TSP-1; angiotatin;
KW endostatin; anticancer; antiangiogenic; cancer; cardiovascular disease;
KW obesity; osteoarthritis; duodenal ulcer; abnormal neovascularisation;
KW wound healing; arteriosclerosis; ischaemic limb; ischaemic myocardium;

KW diabetes mellitus; blood vessel occlusion.
 XX Synthetic.
 OS WO9948924-A1.
 XX 30-SEP-1999.
 XX 23-MAR-1999; 99WO-US006334.
 XX 24-MAR-1998; 98US-00046737.
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 XX (YISS) YISSUM RES & DEV CO.
 XX Ben-Sasson SA;
 XX WPI; 1999-591075/50.
 XX New angiogenic peptide derivatives, used for treating e.g. cancer,
 XX cardiovascular diseases, obesity, osteoarthritis, duodenal ulcers,
 XX abnormal neovascularization and for wound healing.
 XX Disclosure; Page 59; 62pp; English.
 PS The present invention specifically describes peptide derivatives
 XX comprising an angiogenic homology region (AHR) of endostatin. The peptide
 CC derivatives can be used for modulating angiogenesis in humans and
 CC animals. The peptides can be used to treat a wide variety of disease
 CC conditions, including cancer, cardiovascular diseases (e.g.
 CC arteriosclerosis, ischaemic limbs and ischaemic myocardium), obesity,
 CC osteoarthritis, duodenal ulcers, abnormal ocular neovascularisation
 CC associated e.g. with diabetes mellitus, and to promote wound healing or
 CC to stimulate the growth of new blood vessels to bypass, e.g. blood vessel
 CC occlusions. The peptide derivatives can also be used for the production
 CC of antibodies. The multivalent ligands may enable the administration of
 CC lower doses in order to achieve therapeutic efficacy, as compared with a
 CC univalent peptide chain. In addition, they can have long in vivo
 CC lifetimes and good biodistribution when administered orally or
 CC parenterally. The present sequence represents a peptide used in the
 CC exemplification of the present invention
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 112; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 Db 1 GGGGGGGGGGGGGGGGGG 20
 RESULT 4
 AAB03035 100.0%; Score 112; DB 3; Length 20;
 ID AAB03035 standard; peptide; 20 AA.
 XX AAB03035;
 AC
 XX 25-SEP-2000 (first entry)
 DT Flexible linker peptide, SEQ ID NO:116.
 DE MN protein; tumour associated cell adhesion molecule; oncoprotein;
 XX proteoglycan domain; PG domain; carbonic anhydrase; CA domain;
 KW abnormal expression; neoplastic disease; cancer; gene therapy; antibody;
 KW variable domain; flexible linker peptide.
 XX Synthetic.
 OS WO200024913-A2.
 XX 04-MAY-2000.

XX 22-OCT-1999; 99WO-US024879.
 XX 23-OCT-1998; 98US-00177776.
 PR 23-OCT-1998; 98US-00178115.
 XX (FARB) BAYER CORP.
 PA (VIRO-) INST VIROLOGY.
 XX Zavada J, Pastorekova S, Pastorek J;
 PI WPI; 2000-350752/30.
 XX A molecule which specifically binds to a site on MN protein (oncoprotein)
 PT and prevents adhesion of vertebrate cells to the protein, useful for
 PT treating preneoplastic or neoplastic diseases such as cancer.
 XX Claim 22; Page 148; 154pp; English.
 PS The invention relates to the inhibition of cell adhesion mediated by the
 CC MN oncoprotein (also known as the MN/CA IX isoenzyme or the MN/G250
 CC protein). The MN protein is a tumour-associated adhesion molecule which
 CC comprises a proteoglycan-like (PG) domain (AAB03017) which contains the
 CC protein's binding site, and a carbonic anhydrase (CA) domain (AAB03018).
 CC Abnormal expression of the MN protein is associated with tumorigenicity.
 CC The invention encompasses molecules (e.g., proteins and peptides) which
 CC which specifically bind to a site on the MN protein, thereby preventing
 CC adhesion of vertebrate cells to the protein in a cell adhesion assay. It
 CC also encompasses MN proteins or MN protein fragments which can be added
 CC to the extracellular environment to prevent the adhesion of vertebrate
 CC cells to each other. The invention also relates to the identification of
 CC the binding site of the MN protein and to a method of identifying a site
 CC on an MN protein to which cells adhere, comprising testing a series of
 CC overlapping peptides from the protein in a cell adhesion assay. The
 CC invention encompasses a vector comprising an expression control sequence
 CC operatively linked to a nucleic acid encoding the variable domains of a
 CC MN-specific antibody, where the domains are separated by a flexible
 CC linker peptide (AAB03035) and the vector inhibits the growth of a
 CC vertebrate preneoplastic or neoplastic cell that abnormally expresses MN
 CC protein. The invention also encompasses a vector comprising a nucleic
 CC acid encoding a cytotoxic protein or peptide operatively linked to the MN
 CC gene promoter, which inhibits the growth of a vertebrate preneoplastic or
 CC neoplastic cell. Also claimed is a repressor complex that binds to the MN
 CC gene promoter (AAB52473). MN proteins and peptides, MN-binding proteins
 CC and peptides, and expression vectors encoding such proteins and peptides
 CC are useful for treating patients with preneoplastic or neoplastic disease
 CC (e.g., cancers) associated with or characterised by abnormal MN
 CC expression. The present sequence represents a flexible linker peptide
 CC which joins MN-specific antibody variable domains encoded by a vector of
 CC the invention
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 112; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 Db 1 GGGGGGGGGGGGGGGGGG 20
 RESULT 5
 ABP51824
 ID ABP51824 standard; peptide; 20 AA.
 XX ABP51824;
 AC
 XX 07-OCT-2002 (first entry)
 DT Fusion protein related linker peptide C.
 XX Renilla reniformis; green fluorescent protein; GFP; fluorescence;

KW fusion protein; intramolecular dimer fluorescent protein; IDFP; FACS;
 KW multimeric protein; fluorescence activated cell sorting; cell population;
 KW viral vector infection; fluorescence resonance energy transfer; FRET;
 KW transgene expression; protein:protein interaction; humanised; linker.
 XX Synthetic.
 XX OS
 XX PN WO200248174-A2.
 XX PD 20-JUN-2002.
 XX PF 13-DEC-2001; 2001WO-US048690.
 XX PR 15-DEC-2000; 2000US-0256121P.
 XX PA (STRA-) STRATAGENE.
 XX PI Davis RW, Vaillancourt P;
 XX DR WPI; 2002-583482/62.
 XX PT Novel recombinant fusion polypeptide which is fluorescent when excited,
 XX comprising first polypeptide peptide-bonded to second polypeptide, the
 XX polypeptides being found in nature as monomers of a multimeric protein.
 XX PS Disclosure; Fig 4; 59pp; English.
 XX CC The present invention describes a recombinant fusion polypeptide
 CC (intramolecular dimer fluorescent protein (IDFP)) (I) comprising a first
 CC polypeptide peptide-bonded to a second polypeptide, where the first and
 CC second polypeptides are found in nature as monomers of a multimeric
 CC protein and the first and second polypeptides are not fluorescent donor
 CC and acceptor to each other, and (I) is fluorescent when excited. (I) is
 CC useful for monitoring the interaction of two polypeptides of interest
 CC which involves contacting a first polypeptide and the second polypeptide.
 CC (I) is also useful for screening for a compound that modulates the
 CC interaction of first and second member of specific binding pair which
 CC involves contacting a first polypeptide and a second polypeptide in the
 CC presence and absence of candidate modulator. (I) can be used as reporter
 CC genes to monitor the activity of promoter sequences to investigate the
 CC cellular localisation of fusion proteins, to mark cellular proteins for
 CC fluorescence activated cell sorting (FACS) analyses of cell population,
 CC to monitor viral vector infection, to monitor transgene expression in
 CC vivo or in culture, and to monitor protein:protein interactions both in
 CC vivo and in vitro. (I) is suitable for applications that rely on
 CC fluorescence resonance energy transfer (FRET). The present sequence
 CC represents a linker peptide which is given in the exemplification of the
 CC present invention
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 112; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGGGS 20
 | | | | | | | | | | | | | | | | | |
 Db 1 GGGGGGGGGGGGGGGGGGGS 20
 | | | | | | | | | | | | | | | | | |
 RESULT 6
 ABG78248
 ID ABG78248 standard; peptide; 20 AA.
 XX AC
 XX ABG78248;
 XX DT 15-NOV-2002 (first entry)
 XX DE Human Fv molecule hypervariable region related peptide #123.
 XX KW Human; Fv molecule; hypervariable region; single chain Fv; cytostatic;
 KW disulfide Fv; dsFv; scFv; cancer; carcinoma; sarcoma; leukaemia; adenoma;
 KW lymphoma; myeloma; blastoma; seminoma; melanoma; acute myeloid leukaemia.

XX Homo sapiens.
 XX OS
 XX PN WO200259264-A2.
 XX PD 01-AUG-2002.
 XX PF 31-DEC-2001; 2001WO-US049440.
 XX PR 29-DEC-2000; 2000US-00751181.
 XX PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
 XX PI Hagai Y, Lazarovits J, Guy R, Lipschitz O, Szanton E, Levanon A;
 XX PI Plaksin D, Peretz T;
 XX DR WPI; 2002-619166/66.
 XX PT Novel peptide/polypeptide for cancer therapy has Fv molecule, construct
 XX or fragment, or construct of fragment with enhanced binding
 XX characteristics so as to selectively bind target cell in favor of other
 XX cells.
 XX PS Claim 6; Page 211; 232pp; English.
 XX CC The invention relates to a peptide or polypeptide comprising an Fv
 CC molecule, a construct or fragments or a construct of a fragment with
 CC enhanced binding characteristics which selectively and/or specifically
 CC binds to a target cell in favour of other cells, where binding is
 CC primarily determined by a first hypervariable region and Fv is a single
 CC chain Fv (scFv) or a disulfide Fv (dsFv). The peptide, optionally in
 CC association with or attached, coupled, combined, linked or fused to a
 CC pharmaceutical agent, is useful in the manufacture of a medicament, where
 CC the medicament has activity against a diseased cell, preferably a cancer
 CC cell (selected from carcinoma, sarcoma, leukaemia, adenoma, lymphoma,
 CC myeloma, blastoma, seminoma, and melanoma, where the leukaemia cell is an
 CC acute myeloid leukaemia cell). The peptide is also useful for preparing a
 CC composition for use in inhibiting the growth of a diseased or cancer
 CC cell. This sequence represents a human Fv molecule hypervariable region
 CC related peptide of the invention
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 112; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGGGS 20
 | | | | | | | | | | | | | | | | | |
 Db 1 GGGGGGGGGGGGGGGGGGGS 20
 | | | | | | | | | | | | | | | | | |
 RESULT 7
 AAU75147
 ID AAU75147 standard; peptide; 20 AA.
 XX AC
 XX AAU75147;
 XX DT 23-APR-2002 (first entry)
 XX DE Linker peptide used in human antibody invention.
 XX KW Human antibody; immunoglobulin; Ig; variable heavy-chain; VH;
 KW variable light-chain; VL.
 XX OS Synthetic.
 XX PN WO200200729-A2.
 XX PD 03-JAN-2002.
 XX PF 25-JUN-2001; 2001WO-US020542.

PR 23-JUN-2000; 2000US-00602373.
 PR 23-JUN-2000; 2000US-00602972.
 PR 23-JUN-2000; 2000US-00603658.
 PR 23-JUN-2000; 2000US-00603663.
 XX (GENE-) GENETASTIX CORP.
 XX
 XX
 FI Zhu L, Hua SB;
 XX
 XX
 DR WPI; 2002-090521/12.
 DR N-PSDB; ABK32976.
 XX
 XX
 PT Screening libraries of tester proteins against protein, peptide or
 PT nucleic acid target(s) using a two-hybrid method in yeast, useful for
 PT generating recombinant human antibodies and screening for their affinity
 PT binding with target antigens.
 XX
 XX
 PS Example 1; Page 108; 251pp; English.
 XX
 CC The present invention relates to compositions and methods for high
 CC throughput generation and screening of a human antibody or immunoglobulin
 CC (Ig) library in yeast. The method comprises expressing a library of
 CC tester fusion proteins in yeast cells, each tester fusion protein
 CC comprising either an activation domain or a DNA binding domain of a
 CC transcription activator and a tester protein having a large diversity
 CC within the library. The tester protein comprises a first polypeptide
 CC subunit (e.g. human variable heavy-chain, VH) whose sequence varies
 CC within the library, a second polypeptide subunit (e.g. human variable
 CC light-chain, VL) whose sequence varies within the library independently
 CC of the first polypeptide, and a linker peptide which links the first and
 CC second polypeptide subunits. The method is useful for generating
 CC recombinant human antibodies and screening for their affinity binding
 CC with target antigens. The present sequence represents a linker peptide as
 CC described in the methods of the present invention
 XX
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 112; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGGGS 20
 DB 1 GGGGGGGGGGGGGGGGGGGS 20
 RESULT 8
 AAE21140
 ID AAE21140 standard; peptide; 20 AA.
 AC AAE21140;
 XX
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Linker peptide, to join zinc finger domain to C. elegans transposon Tc1.
 XX
 XX
 KW DNA binding protein; zinc finger domain; zinc finger protein; ZFP;
 KW viral replication; gene expression; virucide; transposon; Tc1.
 XX
 XX
 OS Synthetic.
 XX
 FN WO200208286-A2.
 XX
 XX
 PD 31-JAN-2002.
 XX
 XX
 PF 19-JUL-2001; 2001WO-EP008367.
 XX
 XX
 PR 21-JUL-2000; 2000US-0220060P.
 XX
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 XX
 PI Sera T;
 XX

DR WPI; 2002-172000/22.
 XX
 PT New Zinc Finger Protein (ZFP) comprising three essential domains useful
 PT for diagnosing diseases associated with abnormal genomic structure.
 XX
 PS Disclosure; Page 43; 143pp; English.
 XX
 CC The present invention relates to novel DNA binding proteins comprising
 CC zinc finger domains in which two histidine and two cysteine residues
 CC coordinate a central zinc ion. The invention particularly relates to the
 CC identification of a context-independent recognition code to design zinc
 CC finger domains. The invention also relates to zinc finger proteins (ZFP)
 CC designed using this recognition code. The ZFPs are useful for altering
 CC genomic structure, inhibiting viral replication (where viral replication
 CC is inhibited for plant virus, an animal virus or a human virus),
 CC modulating gene expression, detecting an altered zinc finger recognition
 CC sequence and diagnosing disease associated with abnormal genomic
 CC structure. The present sequence is a linker peptide which is used to join
 CC zinc finger domain of the invention to C. elegans transposon Tc1
 XX
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 112; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGGGS 20
 DB 1 GGGGGGGGGGGGGGGGGGGS 20
 RESULT 9
 ABG91939
 ID ABG91939 standard; peptide; 20 AA.
 XX
 XX
 AC ABG91939;
 XX
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE Human antibody fragment #123.
 XX
 XX
 KW Human; antibody; epitope; cancer; tumour; cell rolling; inflammation;
 KW metastasis; hypervariable region; autoimmune disease; thrombosis;
 KW restenosis; leukaemia; inflammatory disease; cardiovascular disease;
 KW myocardial infarction; retinopathic disease; abnormal platelet function;
 KW sulphated tyrosine-dependent protein-protein interaction.
 XX
 XX
 OS Homo sapiens.
 XX
 FN WO200253700-A2.
 XX
 XX
 PD 11-JUL-2002.
 XX
 XX
 PF 31-DEC-2001; 2001WO-US049442.
 XX
 XX
 PR 29-DEC-2000; 2000US-00751181.
 PR 29-DEC-2000; 2000US-0258948P.
 XX
 XX
 PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
 XX
 XX
 PI Lazarovits J, Hagai Y, Plaksin D, Vogel T, Nimrod A, Mar-Haim H;
 PI Szanthon E, Richter T, Amit B, Kooperman L, Peretz T, Levanon A;
 XX
 XX
 DR WPI; 2002-674776/72.
 XX
 XX
 PT Novel isolated epitope present on cancer cells and important in
 PT physiological phenomena such as cell rolling, metastasis and
 PT inflammation, for treating autoimmune, inflammatory or cardiovascular
 PT diseases, and cancer.
 XX
 XX
 PS Disclosure; Page 287; 0pp; English.
 XX
 XX
 CC The invention relates to an isolated epitope present on cancer cells and

CC important in physiological phenomena such as cell rolling, metastasis and
 CC inflammation, where the epitope is capable of being bound by an antibody,
 CC its antigen-binding fragment or its complex comprising at least one
 CC antibody or its binding fragment having a first hypervariable region. The
 CC epitopes are useful for inhibiting cell rolling, inflammation, autoimmune
 CC disease, thrombosis, restenosis, metastasis, growth and/or replication of
 CC tumour or leukaemia cells, increase in number of tumour or leukaemia
 CC cells in a patient, cell-cell, cell-matrix, platelet-matrix, platelet-
 CC platelet and/or cell-platelet adhesion or aggregation, for increasing
 CC mortality of tumour or leukaemia cells, for increasing the susceptibility
 CC of diseased cells to damage by anti-disease, anti-cancer or anti-
 CC leukaemia agents, or for decreasing the number of tumour or leukaemia
 CC cells in a patient, or in the manufacture of a medicament for the above
 CC mentioned purposes. The epitopes are useful for diagnosing and treating
 CC diseases such as cancer, leukaemia, autoimmune diseases, inflammatory
 CC diseases, cardiovascular diseases such as myocardial infarction,
 CC retinopathic diseases and other diseases mediated by abnormal platelet
 CC function and diseases caused by sulphated tyrosine-dependent protein-
 CC protein interactions. This sequence represents a human antibody fragment
 CC of the invention
 CC
 CC SQ Sequence 20 AA;

Query Match 100.0%; Score 112; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGGGGGGGGGGGGGGGGG 20
 Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 10
 ABR57056
 ID ABR57056 standard; peptide; 20 AA.
 AC ABR57056;
 XX
 XX 05-AUG-2003 (first entry)
 DT Peptide linker LL.
 DE Multimeric single chain tandem Fv-antibody; antibacterial; virucide;
 XX cytostatic; cytokine antagonist; diagnosis; viral disease; prion disease;
 KW bacterial disease; tumoural disease; agglutination; red blood cell;
 KW immune system; tumour cell; cytokine; cytotoxic; linker.
 XX
 XX Synthetic.
 OS
 XX
 XX EP1293514-A1.
 XX
 XX 19-MAR-2003.
 XX
 XX 14-SEP-2001; 2001EP-00122104.
 XX
 XX 14-SEP-2001; 2001EP-00122104.
 XX
 XX (AFFI-) AFFIMED THERAPEUTICS AG.

XX Le Gall F, Kipriyanov S, Reusch U, Moldenhauer G, Little M;
 XX WPI; 2003-395536/38.
 XX
 XX New multimeric Fv-antibody having monomers forming antigen-binding units
 PT and sites, useful for the preparation of a pharmaceutical composition for
 PT the treatment of a viral, bacterial, tumoral or prion diseases.
 XX
 XX Example 1; Page 6; 29pp; English.

XX The present invention describes a multimeric Fv-antibody (I) comprising:
 CC (a) monomers comprising at least 4 variable domains of which two
 CC neighbouring domains of one monomer form an antigen-binding VH-VL or VL-
 CC VH scFv unit; and/or (b) at least 2 variable domains of a monomer that

CC are non-covalently bound to 2 variable domains of another monomer
 CC resulting in the formation of at least 2 additional antigen binding sites
 CC to form multimerisation motif. Also described is a process for the
 CC preparation of (I) comprising ligating DNA sequences encoding the peptide
 CC linkers with the DNA sequences encoding the variable domains such that
 CC the peptide linkers connect the variable domains resulting in the
 CC formation of a DNA sequence encoding a monomer of the multivalent
 CC multimeric Fv-antibody, and expressing the DNA sequences encoding the
 CC various monomers in an expression system. (I) has antibacterial, virucide
 CC and cytostatic activities, and can be used as a cytokine antagonist. The
 CC multimeric Fv-antibody is useful for diagnosis. The antibody can also be
 CC used for the preparation of a pharmaceutical composition for the
 CC treatment of a viral, bacterial, tumoural or prion disease, the
 CC agglutination of red blood cells, linking cytotoxic cells of the immune
 CC system to tumour cells, or linking activating cytokines, cytotoxic
 CC substances or a protease to a target cell. The present sequence
 CC represents a linker which is used in the exemplification of the present
 CC invention
 CC
 CC SQ Sequence 20 AA;

Query Match 100.0%; Score 112; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGGGGGGGGGGGGGGGGG 20
 Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 11
 AAE37194
 ID AAE37194 standard; peptide; 20 AA.
 AC AAE37194;
 XX
 XX 07-AUG-2003 (first entry)
 DT Linker peptide #2.
 DE CD4-inducible epitope; Human immunodeficiency virus; HIV; gene therapy;
 KW HIV infection; envelope glycoprotein; Env; vaccine.
 XX
 XX Unidentified.
 OS
 XX WO2003033666-A2.
 XX
 XX 24-APR-2003.
 XX
 XX 16-OCT-2002; 2002WO-US033165.
 XX
 XX 16-OCT-2001; 2001US-0329709P.
 XX
 XX (SCRI) SCRIPPS RES INST.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Dimitrov DS, Mouldard M, Xiao X, Shu Y, Phogat SK, Zhang M;
 XX Burton D;
 XX WPI; 2003-393518/37.
 XX
 XX New isolated antibody or antibody fragment specifically binds a CD4-
 PT inducible epitope on Human Immunodeficiency Virus (HIV) Env proteins,
 PT useful for preparing a composition for treating or preventing HIV
 PT infection.
 XX
 XX Example 3; Page 67; 69pp; English.
 XX
 XX The invention relates to antibodies or antibody fragments specifically
 CC binding to CD4-inducible epitope on Human immunodeficiency virus (HIV)
 CC envelope glycoprotein (Env) proteins. The antibody is useful for
 CC preparing a composition for treating or preventing HIV infection. The
 CC invention is used to prepare vaccines and is used in gene therapy. The

CC present sequence is a linker peptide used in the exemplification of the
 CC invention
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 112; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
 Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 12
 ABP60059
 ID ABP60059 standard; peptide; 20 AA.
 AC ABP60059;
 XX
 DT 21-MAR-2003 (first entry)
 XX
 DE Flexible linker peptide # SEQ ID 9.
 XX
 DE Clinical diagnosis; food hygiene; environmental monitoring; measure;
 KW energy transfer.
 XX
 OS Synthetic.
 XX
 PN WO200288733-A1.
 XX
 PD 07-NOV-2002.
 XX
 PF 23-APR-2002; 2002WO-JP004022.
 XX
 PR 23-APR-2001; 2001JP-00123866.
 XX
 PA (EIKE) EIKEN KAGAKU KK.
 XX
 PA (NAGA) NAGAWONE T.
 XX
 PI Nagamune T, Ueda H, Ohiro Y, Shibata N;
 XX
 DR WPI; 2003-111910/10.
 XX
 XX Analysis method using reporter (label) intermolecular interaction based
 PT on energy transfer phenomenon, useful in clinical diagnosis, food hygiene
 PT and environmental monitoring.
 XX
 PS Example 2; Page 14; 34pp; Japanese.
 XX
 CC The invention relates to a method for the analysis of a substance and
 CC comprises labelling substances (A and B) with binding affinity to the
 CC substance to be measured (X) with reagent-labelled reporters (R1 and R2),
 CC labelling substances (C and D) with the reagents, and mixing with a
 CC sample to produce a reaction composite. Energy transfer due to
 CC interaction between the 2 reporters is then measured. The method is used
 CC for measuring the concentration of a substance, e.g. a protein, peptide,
 CC antigen, antibody, lectin, lectin-binding sugar, tumour marker, cytokine,
 CC cytokine receptor, hormone, hormone receptor, cell-adhesion factor,
 CC ligand of cell-adhesion factor, nucleic acid, sugar chain and lipid, a
 CC cell, an intracellular organelle or a low-molecular compound. The method
 CC of the invention is useful in clinical diagnosis, food hygiene and
 CC environmental monitoring. The method is highly accurate and sensitive,
 CC with efficient energy transfer to lower background interference in a
 CC homogeneous system. It is easy to operate and can be applied in situ. The
 CC current sequence represents a flexible linker peptide created in an
 CC example from the invention using 2 primers for site-directed mutagenesis
 CC as given in records ABZ56902-3
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 112; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
 Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 13
 AAG79954
 ID AAG79954 standard; peptide; 20 AA.
 AC AAG79954;
 XX
 DT 21-MAY-2003 (first entry)
 XX
 DE VH-VL linker sequence in scFv master gene of the HuCAL library.
 XX
 KW Variable; heavy; VH; domain; VH1a; subclass; VH1b; VH2; VH4; VH5; VH6;
 KW light; VL; VLkappa2; VLlambdal; linker; scFv; HuCAL; expression vector;
 KW pBS13; antibody.
 XX
 OS Synthetic.
 XX
 PN WO2003008451-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 19-JUL-2002; 2002WO-EP008094.
 XX
 PR 19-JUL-2001; 2001EP-00116756.
 XX
 PA (UVZU-) UNIV ZUERICH.
 XX
 PI Ewert S, Huber T, Honegger A, Plueckthun A;
 XX
 DR WPI; 2003-229560/22.
 XX
 XX New isolated polypeptides containing a variable heavy domain or variable
 PT light domain of a particular subclass, which comprise modified framework
 PT residues, so that their biophysical properties are improved.
 XX
 PS Example 1; Page 29; 123pp; English.
 XX
 CC The sequences given in AAG79954-65 are peptides which were used in the
 CC construction of the isolated polypeptides of the invention. Peptide (I)
 CC comprises a variable heavy (VH) domain belonging to VH1a subclass, VH1b
 CC subclass, VH2 subclass, VH4 subclass, VH5 subclass or VH6 subclass which
 CC comprises modified framework residues, and peptide (II) comprises a
 CC variable light (VL) domain belonging to VLkappa2 subclass or VLlambdal
 CC subclass which comprises modified framework residues. This sequence
 CC represents the linker between the VH and VL regions of the scFv master
 CC gene of the HuCAL library in the expression vector pBS13. Nucleic acid
 CC encoding (I) or (II) are useful for producing a VH or VL domain or an
 CC antibody or a functional fragment which involves expressing the
 CC respective nucleic acids. These nucleic acids are useful for obtaining a
 CC polypeptide, which involves expressing the nucleic acid sequence. They
 CC are also useful for constructing a library of antibodies or their
 CC functional fragments, which involves obtaining at least one nucleic acid
 CC and diversifying the obtained nucleic acid sequence to generate a
 CC population of diversified nucleic acid sequences, where the diversified
 CC nucleic acid sequences can be expressed for generating and screening of
 CC antibody libraries comprising diversified VH or VL domains. (I) and (II)
 CC have improved biophysical properties. The properties of any given heavy
 CC or light chain variable region can be improved while maintaining the
 CC binding activity of the variable domains
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 112; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20

```

Db      1 GGGSGGGSGGGSGGGGS 20
|||||
RESULT 14
ABR62592
ID   ABR62592 standard; peptide; 20 AA.
XX
AC   ABR62592;
XX
DT   06-NOV-2003 (first entry)
XX
DE   Peptide linker.
XX
KW   CD7; antibody; TH-69; scFv; exotoxin A; immunotoxin; cytostatic;
KW   immunosuppressive; vaccine; linker.
XX
OS   Synthetic.
XX
PN   WO2003051926-A2.
XX
PD   26-JUN-2003.
XX
PF   11-DEC-2002; 2002WO-EP014064.
XX
PR   14-DEC-2001; 2001US-0339422P.
XX
PA   (UYER-) UNIV ERLANGEN-NUERNBURG.
XX
PI   Fey GHM, Gramatzki M, Peipp M;
XX
DR   WPI; 2003-523519/49.
XX
DR   N-PSDB; ACF05483.
XX
XX
XX   New fusion protein comprising functionally linked components of an anti-
PT   CD7 antibody or its fragment and an immunotoxin, useful for treating CD7
PT   positive acute T-cell and/or myeloid leukemias, or Graft-versus-host
PT   disease.
XX
XX   Claim 8; Page 7; 55pp; English.
XX
XX   The present sequence is that of a peptide linker, which was used in novel
CC   fusion proteins of the invention (see ABR62590 and ABR62591), joining an
CC   scFv fragment of anti-CD7 monoclonal antibody TH-69 and domains II and
CC   III of Pseudomonas exotoxin A (ETA). These novel fusion proteins bind to
CC   CD7-positive T-lymphoid cells and kill them by the induction of
CC   apoptosis. Use of the peptide linker avoids some of the problems
CC   encountered with earlier CD7-directed immunotoxins generated by chemical
CC   coupling of a toxin to an antibody. The fusion protein is used in the
CC   treatment of disorders involved a hyperproliferation of CD7-positive
CC   cells, especially acute T-cell and/or myeloid leukaemia, and also in the
CC   treatment or prophylaxis of Graft-versus-host disease, and to condition
CC   an animal (including a human) to be transplanted with donor cells, tissue
CC   or organ
XX
XX   Sequence 20 AA;
XX
Query Match      100.0%; Score 112; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGSGGGSGGGSGGGGS 20
|||||
Db      1 GGGSGGGSGGGSGGGGS 20
|||||

RESULT 16
ADC34709
ID   ADC34709 standard; peptide; 20 AA.
XX
AC   ADC34709;
XX
DT   18-DEC-2003 (first entry)
XX
DE   (Ser4Gly)4 protease substrate, PFPS12.
XX
KW   Gene delivery; fibroblast growth factor; FGF; FGF receptor; gene therapy;
KW   hyperproliferative disorder; neoplastic disease; ligand;
KW   nucleic acid binding domain; NABD; cytocide; therapeutic; prodrug;
KW   ribozyme; antisense; ribosome inactivating protein; saporin; gelonin;
KW   diphtheria toxin; elongation factor 2; HSV thymidine kinase;
KW   cytosine deaminase; protease; PFPS12.
XX
OS   Synthetic.
XX
PN   US6503886-B1.

```

```

Db      1 GGGSGGGSGGGSGGGGS 20
|||||
RESULT 14
ABR62592
ID   ABR62592 standard; peptide; 20 AA.
XX
AC   ABR62592;
XX
DT   06-NOV-2003 (first entry)
XX
DE   Peptide linker.
XX
KW   CD7; antibody; TH-69; scFv; exotoxin A; immunotoxin; cytostatic;
KW   immunosuppressive; vaccine; linker.
XX
OS   Synthetic.
XX
PN   WO2003051926-A2.
XX
PD   26-JUN-2003.
XX
PF   11-DEC-2002; 2002WO-EP014064.
XX
PR   14-DEC-2001; 2001US-0339422P.
XX
PA   (UYER-) UNIV ERLANGEN-NUERNBURG.
XX
PI   Fey GHM, Gramatzki M, Peipp M;
XX
DR   WPI; 2003-523519/49.
XX
DR   N-PSDB; ACF05483.
XX
XX
XX   New fusion protein comprising functionally linked components of an anti-
PT   CD7 antibody or its fragment and an immunotoxin, useful for treating CD7
PT   positive acute T-cell and/or myeloid leukemias, or Graft-versus-host
PT   disease.
XX
XX   Claim 8; Page 7; 55pp; English.
XX
XX   The present sequence is that of a peptide linker, which was used in novel
CC   fusion proteins of the invention (see ABR62590 and ABR62591), joining an
CC   scFv fragment of anti-CD7 monoclonal antibody TH-69 and domains II and
CC   III of Pseudomonas exotoxin A (ETA). These novel fusion proteins bind to
CC   CD7-positive T-lymphoid cells and kill them by the induction of
CC   apoptosis. Use of the peptide linker avoids some of the problems
CC   encountered with earlier CD7-directed immunotoxins generated by chemical
CC   coupling of a toxin to an antibody. The fusion protein is used in the
CC   treatment of disorders involved a hyperproliferation of CD7-positive
CC   cells, especially acute T-cell and/or myeloid leukaemia, and also in the
CC   treatment or prophylaxis of Graft-versus-host disease, and to condition
CC   an animal (including a human) to be transplanted with donor cells, tissue
CC   or organ
XX
XX   Sequence 20 AA;
XX
Query Match      100.0%; Score 112; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGSGGGSGGGSGGGGS 20
|||||
Db      1 GGGSGGGSGGGSGGGGS 20
|||||

RESULT 15
AAE38680
ID   AAE38680 standard; peptide; 20 AA.
XX
AC   AAE38680;
XX
DT   04-DEC-2003 (first entry)
XX

```

```

XX
DE   Linker peptide used to join a zinc finger domain to C. elegans Tc1.
XX
XX   Artificial transcription factor; DNA binding protein; ATF; ZFP; therapy;
KW   zinc finger protein; crop protection; disease-resistant; transgenic;
KW   transgenic plant.
XX
OS   Unidentified.
XX
PN   WO2003062455-A2.
XX
PD   31-JUL-2003.
XX
PF   23-JAN-2003; 2003WO-US002358.
XX
PR   23-JAN-2002; 2002US-00057408.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Sera T;
XX
DR   WPI; 2003-646071/61.
XX
XX   Preparing an artificial transcription factor (ATF) capable of modulating
PT   expression of a gene by interaction with a target site associated with
PT   the gene, for treating plant disease, comprises preparing a combinatorial
PT   library of ATFs.
XX
XX   Disclosure; Page 67; Opp; English.
XX
XX   The invention relates to a method of preparing artificial transcription
CC   factor (ATF) capable of modulating expression of a gene by interaction
CC   with a target site associated with the gene. The method comprises
CC   preparing a combinatorial library of ATFs, each of the ATFs comprising a
CC   DNA-binding domain and a transcriptional regulatory domain. The invention
CC   also relates to DNA binding proteins comprising zinc finger domains and
CC   particularly to the identification of a context-independent recognition
CC   code to zinc finger domains. The methods are useful for treating disease
CC   in a plant, for crop protection and for producing genetically transformed
CC   disease-resistant plants. The present sequence is a linker peptide used
CC   to join a zinc finger domain to C. elegans Tc1 used in the invention
XX
XX   Sequence 20 AA;
XX
Query Match      100.0%; Score 112; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGGSGGGSGGGSGGGGS 20
|||||
Db      1 GGGSGGGSGGGSGGGGS 20
|||||

RESULT 16
ADC34709
ID   ADC34709 standard; peptide; 20 AA.
XX
AC   ADC34709;
XX
DT   18-DEC-2003 (first entry)
XX
DE   (Ser4Gly)4 protease substrate, PFPS12.
XX
KW   Gene delivery; fibroblast growth factor; FGF; FGF receptor; gene therapy;
KW   hyperproliferative disorder; neoplastic disease; ligand;
KW   nucleic acid binding domain; NABD; cytocide; therapeutic; prodrug;
KW   ribozyme; antisense; ribosome inactivating protein; saporin; gelonin;
KW   diphtheria toxin; elongation factor 2; HSV thymidine kinase;
KW   cytosine deaminase; protease; PFPS12.
XX
OS   Synthetic.
XX
PN   US6503886-B1.

```

```
XX 07-JAN-2003.
XX 24-NOV-1999; 99US-00449249.
XX 15-MAR-1994; 94US-00213446.
XX 15-MAR-1994; 94US-00213447.
XX 29-AUG-1994; 94US-00297961.
XX 13-SEP-1994; 94US-00305771.
XX 16-MAY-1995; 95US-00441979.
XX 24-SEP-1996; 96US-00718904.
XX (SELE-) SELECTIVE GENETICS INC.
XX Baird JA, Chandler LA, Sosnowski BA;
XX WPI; 2003-361554/34.
XX N-PSDB; ADC34615.
XX Gene delivery composition for therapeutic treatments, comprises
XX polypeptide that binds to fibroblast growth factor receptor-nucleic acid
XX molecule.
XX Example 7; Col 68; 130pp; English.
XX The invention discloses a gene delivery composition comprising a
XX polypeptide that binds to a fibroblast growth factor (FGF) receptor-
XX nucleic acid molecule. The nucleic acid molecule is chemically conjugated
XX or fused to the polypeptide that binds to an FGF receptor. The delivery
XX composition binds to an FGF receptor and is internalised specifically in
XX cells bearing the FGF receptor. The invention relates to the treatment
XX (e.g. gene therapy) of diseases (e.g. hyperproliferative disorders and
XX neoplastic diseases) and, more specifically, to the preparation and use
XX of complexes containing receptor-binding internalised ligands, nucleic
XX acid binding domain (NABD) and cytotoxic-encoding agents to alter the
XX function, gene expression or viability of a cell in a therapeutic manner.
XX The nucleic acid molecule is a prodrug-encoding agent, a ribozyme or an
XX antisense molecule. The cytotoxic-encoding agent encodes a ribosome
XX inactivating protein, preferably a saporin, gelonin or diphtheria toxin.
XX It also encodes its elongation factor 2 and further comprises a tissue-
XX specific promoter operably linked to the molecule. The prodrug-encoding
XX agent encodes HSV thymidine kinase or cytosine deaminase. The polypeptide
XX that binds to an FGF receptor can be FGF-1 polypeptide, FGF-2
XX polypeptide, FGF-3 polypeptide, FGF-4 polypeptide, FGF-5 polypeptide, FGF
XX -6 polypeptide, FGF-7 polypeptide, FGF-8 polypeptide or FGF-9
XX polypeptide. The invention composition has increased specificity and
XX delivers higher amounts of nucleic acids to targeted cells. The sequence
XX presented is a protease substrate.
XX Sequence 20 AA;
XX Query Match 100.0%; Score 112; DB 7; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-05;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 GGGGGGGGGGGGGGGGGGGG 20
XX |||||
XX DB 1 GGGGGGGGGGGGGGGGGGGG 20
XX |||||
XX RESULT 18
XX ADN97318
XX ID ADN97318 standard; peptide; 20 AA.
XX AC ADN97318;
XX DT 01-JUL-2004 (first entry)
XX DE Linker peptide SEQ ID NO:4.
XX KW linker; antibody variable region; scDb library; immune diagnosis;
XX gene therapy; immunoassay; bispecific Db library; scFv library.
XX OS Synthetic.
XX XX WO2003087163-A1.
XX PN 23-OCT-2003.
XX PD
XX PF 15-APR-2003; 2003WO-JP004773.
XX PR 15-APR-2002; 2002JP-00112369.
XX PA (CHUS ) CHUGAI SEIYAKU KK.
KW cancer; transplant rejection; infectious disease; inflammation; linker.
XX Synthetic.
XX US2003152913-A1.
XX 14-AUG-2003.
XX 08-FEB-2002; 2002US-00072301.
XX 08-FEB-2002; 2002US-00072301.
XX (HUAS/) HUA S.
XX (PAUL/) PAULING M H.
XX (ZHUL/) ZHU L.
XX Hua S, Pauling MH, Zhu L;
XX WPI; 2003-897702/82.
XX New human monoclonal antibody binding loop 6 of human chemokine receptor,
XX CCR5, for high throughput screening of an antibody library, and for
XX diagnosing, preventing and/or treating human immunodeficiency virus
XX infection.
XX Example; SEQ ID NO 4; 64pp; English.
XX The invention describes an antibody (I) that binds to loop 6 of human
XX CCR5 (chemokine receptor) and inhibits infection of immunodeficiency
XX virus of human cells. The methods and compositions of the present
XX invention are useful in high throughput screening of an antibody library
XX against a wide variety of target proteins, and for diagnosing, preventing
XX and/or treating HIV infection, autoimmune diseases, cancer, transplant
XX rejection, infectious diseases and inflammation. This is the amino acid
XX sequence of an anti-human CC chemokine receptor 5 (CCR5)-antibody
XX associated linker peptide used to link the heavy chain variable regions
XX and light chain variable regions.
XX Sequence 20 AA;
XX Query Match 100.0%; Score 112; DB 7; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-05;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 GGGGGGGGGGGGGGGGGGGG 20
XX |||||
XX DB 1 GGGGGGGGGGGGGGGGGGGG 20
XX |||||
XX RESULT 18
XX ADN97318
XX ID ADN97318 standard; peptide; 20 AA.
XX AC ADN97318;
XX DT 01-JUL-2004 (first entry)
XX DE Linker peptide SEQ ID NO:4.
XX KW linker; antibody variable region; scDb library; immune diagnosis;
XX gene therapy; immunoassay; bispecific Db library; scFv library.
XX OS Synthetic.
XX XX WO2003087163-A1.
XX PN 23-OCT-2003.
XX PD
XX PF 15-APR-2003; 2003WO-JP004773.
XX PR 15-APR-2002; 2002JP-00112369.
XX PA (CHUS ) CHUGAI SEIYAKU KK.
```

XX Kojima T;
 XX WPI; 2003-833711/77.
 XX N-PSDB; ADN97317.
 XX Construction of scDb (single-chain diabody) library by disposing
 PT restriction enzyme sites appropriately for antigen-encoding regions, for
 PT use in immune diagnosis, gene therapy and immunoassay.
 XX
 XX Disclosure; SEQ ID NO 4; 36pp; Japanese.
 XX
 CC The invention relates to a novel gene encoding 2 antibody variable
 CC regions with a linker containing restriction enzyme sites connected to
 CC one and the other side of such antibody variable regions. The constructed
 CC scDb libraries of the invention are useful in immune diagnosis, gene
 CC therapy and immunoassay. With this method, a bispecific Db library can be
 CC efficiently and collectively constructed from an scFv library without
 CC resorting to any troublesome procedures. The present sequence represents
 CC a linker peptide used in the invention.
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 100.0%; Score 112; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 DB 1 GGGGGGGGGGGGGGGGGG 20
 RESULT 19
 ADN97316
 ID ADN97316 standard; peptide; 20 AA.
 XX
 AC ADN97316;
 DT 01-JUL-2004 (first entry)
 XX
 DE Linker peptide SEQ ID NO:2.
 XX
 KW linker; antibody variable region; scDb library; immune diagnosis;
 KW gene therapy; immunoassay; bispecific Db library; scFv library.
 XX
 OS Synthetic.
 XX
 XX WO2003087163-A1.
 XX
 XX 23-OCT-2003.
 PD
 XX 15-APR-2003; 2003WO-JP004773.
 XX
 PF 15-APR-2002; 2002JP-00112369.
 XX
 PR (CHUS) CHUGAI SEIYAKU KK.
 XX
 PA Kojima T;
 XX
 PI WPI; 2003-833711/77.
 XX
 DR N-PSDB; ADN97315.
 DR
 XX Construction of scDb (single-chain diabody) library by disposing
 PT restriction enzyme sites appropriately for antigen-encoding regions, for
 PT use in immune diagnosis, gene therapy and immunoassay.
 XX
 XX Disclosure; SEQ ID NO 2; 36pp; Japanese.
 XX
 CC The invention relates to a novel gene encoding 2 antibody variable
 CC regions with a linker containing restriction enzyme sites connected to
 CC one and the other side of such antibody variable regions. The constructed
 CC scDb libraries of the invention are useful in immune diagnosis, gene
 CC therapy and immunoassay. With this method, a bispecific Db library can be

CC efficiently and collectively constructed from an scFv library without
 CC resorting to any troublesome procedures. The present sequence represents
 CC a linker peptide used in the invention.
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 100.0%; Score 112; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 DB 1 GGGGGGGGGGGGGGGGGG 20
 RESULT 20
 ADG42781
 ID ADG42781 standard; peptide; 20 AA.
 XX
 AC ADG42781;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Peptide linker.
 XX
 KW HIV infection; HIV type 1; HIV type 2; cancer; breast; prostate; liver;
 KW renal; lung; skin; ovarian; cervical; brain; thyroid; stomach; colon;
 KW lymphoma; leukaemia; pancreas.
 XX
 OS Synthetic.
 XX
 XX US2003206909-A1.
 PN
 PD 06-NOV-2003.
 XX
 XX 07-FEB-2003; 2003US-00360828.
 PF
 XX 08-FEB-2002; 2002US-00071866.
 PR
 PR 08-FEB-2002; 2002US-00072301.
 PR
 PR 25-APR-2002; 2002US-00133978.
 XX
 XX (HUAS/) HUA S.
 PA
 PA (PAUL/) PAULING M H.
 PA (ZHUL/) ZHU L.
 XX
 PI Hua S, Pauling MH, Zhu L;
 XX
 XX WPI; 2004-051479/05.
 XX
 XX Antibody binding loop of human chemokine receptor useful for the
 PT treatment of HIV infection and cancer.
 XX
 XX Disclosure; SEQ ID NO 4; 74pp; English.
 XX
 CC The invention relates to an antibody that binds to loop 6 of human
 CC chemokine receptor (CXCR4). The antibody is useful in the treatment or
 CC prevention of HIV infection (e.g. HIV type 1 and HIV type 2) and cancer
 CC of (e.g. breast, prostate, liver, renal, lung, skin, ovarian, cervical,
 CC brain, thyroid, stomach, colon, lymphoma, leukaemia and pancreas cancer
 CC cells), for screening drugs, for diagnosing disease or condition
 CC associated with interaction with chemokine receptor. The present sequence
 CC is used in the exemplification of the present invention.
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 100.0%; Score 112; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 DB 1 GGGGGGGGGGGGGGGGGG 20

CC sarcoma, breast cancer, colon cancer, prostate cancer and thyroid
 CC cancer), diabetic retinopathy, diabetic nephropathy, other complications
 CC of diabetes, and acromegaly. The present sequence represents a (Gly4Ser)⁴
 CC linker which was used to link the GH mutant GHa to a fragment of the GHR
 CC extracellular domain in the example of the invention. Note: The present
 CC sequence is derived from that shown on page 5.

XX Sequence 20 AA;

Query Match 100.0%; Score 112; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
 |||||
 Db 1 GGGGGGGGGGGGGGGG 20

RESULT 23

AD021851
 ID ADO21851 standard; peptide; 20 AA.

XX AC ADO21851;

XX DT 12-AUG-2004 (first entry)

XX DE Histone acetylation fusion protein reporter-related linker peptide 12.

XX KW fusion protein reporter; histone modification; cytostatic;
 XX KW neuroprotective; muscular; virucide; cancer; neurological;
 XX KW muscular disorders; developmental; cell development; toxin;
 XX KW toxic chemical; virus; linker.

XX OS Synthetic.

XX PN WO200404168-A2.

XX FD 27-MAY-2004.

XX PF 12-NOV-2003; 2003WO-US036059.

XX PR 12-NOV-2002; 2002US-0425578P.

XX PR 05-AUG-2003; 2003US-00634740.

XX PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX PI Ting AY;

XX PS WPI; 2004-420310/39.

XX PT New fusion protein reporter having a core with a histone-modification-
 PT specific binding domain conjugated to a histone polypeptide, useful for
 PT diagnosing or treating cancer, viral infections, neurological and
 PT muscular disorders.

XX PS Disclosure; SEQ ID NO 20; 96pp; English.

XX CC The invention relates to a novel fusion protein reporter comprising a
 CC core having a histone-modification-specific binding domain conjugated to
 CC a histone polypeptide, wherein the core is flanked by donor and acceptor
 CC fluorescent moieties. The molecules of the invention demonstrate
 CC cytosolic, neuroprotective, muscular and virucide activities and may be
 CC useful for the diagnosis and/or treatment of histone modification-
 CC associated disorders, such as cancer, neurological disorders, muscular
 CC disorders, developmental disorders, cell development disorders and
 CC disorders related to exposure to toxins, toxic chemicals or viruses. The
 CC current sequence is that of the histone acetylation fusion protein
 CC reporter-related linker peptide 12 of the invention.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 112; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
 |||||
 Db 1 GGGGGGGGGGGGGGGG 20

RESULT 24

ADS91462

ID ADS91462 standard; peptide; 20 AA.

XX AC ADS91462;

XX DT 02-DEC-2004 (first entry)

XX DE Peptide sequence of spacer, seq ID #84.

XX KW Spacer; multimeric protein; egg-laying; avian; milk producing animal;
 XX KW transgenic; transposase; cancer; immunotherapy; infectious disease.

XX OS Unidentified.

XX PN WO2004067706-A2.

XX PD 12-AUG-2004.

XX PF 24-DEC-2003; 2003WO-US041261.

XX PR 21-JAN-2003; 2003US-0441377P.

XX PR 21-JAN-2003; 2003US-0441381P.

XX PR 21-JAN-2003; 2003US-0441392P.

XX PR 21-JAN-2003; 2003US-0441405P.

XX PR 21-JAN-2003; 2003US-0441447P.

XX PR 21-JAN-2003; 2003US-0441502P.

XX PR 26-JUN-2003; 2003US-00609019.

XX PA (TRAN-) TRANSGENRX INC.

XX PA (LOU) UNIV LOUISIANA STATE & AGRIC & MECH COLL.

XX PI Cooper RK, Fioretti WC, Cadd GG;

XX PS WPI; 2004-580985/56.

XX DR New isolated polynucleotide having two or more genes of interest and pro
 PT nucleotide sequences, for use in pharmaceutical, diagnostic and pro
 PT industrial uses, such as in treating and diagnosing cancer.

XX PS Disclosure; SEQ ID NO 84; 119pp; English.

XX CC The present invention provides a new, effective and efficient method for
 CC producing multimeric proteins in the egg-laying (avian) or milk producing
 CC animals. The multimeric proteins include associated multimeric proteins
 CC (two or more associated polypeptides) and multivalent multimeric proteins
 CC (a single polypeptide is encoded by multiple genes). The expression and
 CC formation of a multimeric protein is achieved by administering a
 CC polynucleotide cassette containing genes of interest (for example
 CC proinsulin, antibodies, cytokines, etc.). The polynucleotide cassettes
 CC are administered to egg-laying or milk producing animals to create the
 CC transgenic animals. Several different signal sequences and promoters can
 CC be used to achieve the deposition of the multimeric protein in the egg or
 CC milk. The transposon based vector has been used for the incorporation of
 CC polynucleotide cassettes into the genomic DNA. The polynucleotide
 CC cassette may additionally contain multiple pro sequences, prepro
 CC sequences, cecropin prepro sequences and cleavage sites. Each gene
 CC encodes a polypeptide which forms part of a multimeric protein. The pro
 CC portion of the sequences has been used to facilitate appropriate
 CC processing, expression and formation of multimeric proteins. These
 CC multimeric proteins can be produced much more efficiently and
 CC economically by this method than earlier conventional methods. This
 CC improved method can produce the multimeric proteins on a large scale. The
 CC transposon-based vectors of the invention produce high integration
 CC frequencies compared to the earlier vectors. The intratesticular
 CC injections with the novel transposon based vector produces 77 % sperm

CC positive roosters (compared to 41% of a previous method). The transposon-
 CC based vectors of the invention include a transposase gene linked to a
 CC first promoter and a coding sequence of a desired protein linked to a
 CC second promoter. The coding sequence for the desired protein and its
 CC promoter are flanked by transposase insertion sequences recognised by the
 CC transposase. The transposon based vectors also include multiple Kozak
 CC sequences (which enhance the expression of the transposase gene),
 CC modified multiple codons at the N-terminal end (to facilitate the
 CC transcription of the transposase gene), an effective polyA sequence (to
 CC further enhance the expression of the transposase gene) and an additional
 CC stop codon (to enhance the termination of transposon synthesis). The
 CC method of the invention can be used to produce antibodies for cancer
 CC immunotherapy, infectious disease and toxic agents. The methods and
 CC compositions of producing multimeric proteins are useful in
 CC pharmaceutical, diagnostic and industrial uses (treatment and diagnosis
 CC of cancer). The polypeptide cassettes also include one or more epitopes
 CC or domains for facilitating purification of a desired protein. The
 CC presented sequence is the spacer peptide, seq ID #84.

XX Sequence 20 AA;

Query Match 100.0%; Score 112; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
 |||||
 Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 25

ADT62591 ADT62591 standard; peptide; 20 AA.

XX AC ADT62591;

DT 16-DEC-2004 (first entry)

XX spacer sequence.

XX Transposon-based vector; transposon; transgenic; vaccine;
 KW immune response; anthrax; botulism; brucellosis; glanders; Q fever;
 KW plague; shigellosis; small pox; tularemia; viral encephalitis;
 KW typhus fever; viral hemorrhagic fever; bird; mammal; egg; milk;
 KW antibacterial; anti viral; antimicrobial; anti-inflammatory;
 KW neuroprotective; transposase; spacer.

XX Unidentified.

OS WO2004067743-A1.

FN 12-AUG-2004.

PD 24-DEC-2003; 2003WO-US041269.

XX 21-JAN-2003; 2003US-0441377P.

PR 21-JAN-2003; 2003US-0441381P.

PR 21-JAN-2003; 2003US-0441392P.

PR 21-JAN-2003; 2003US-0441405P.

PR 21-JAN-2003; 2003US-0441447P.

PR 21-JAN-2003; 2003US-0441502P.

PR 26-JUN-2003; 2003US-00609019.

XX (TRAN-) TRANSGENX INC.

PA (LOU) UNIV LOUISIANA STATE & AGRIC & MECH COLL.

XX Cooper RK, Fioretti WC, Cadd GG;

XX WPI; 2004-581003/56.

XX New polynucleotide (transposon-based vector) cassette comprising gene of
 PT interest operably linked to a cecropin prepro sequence, is useful in
 PT producing vaccines to protect an individual against infectious diseases.

XX Disclosure; SEQ ID NO 84; 155pp; English.

XX The patent discloses novel, effective and efficient method of producing
 CC multimeric proteins, antibodies using transposon-based vectors, in
 CC transgenic individuals. These proteins can be used as a vaccine.
 CC Immunising individuals with such compositions, comprising vaccines is
 CC capable of generating an immune response, to prevent or ameliorate the
 CC severity of a disease. The invention provides polynucleotide cassettes
 CC containing at least one gene of interest and one or more pro
 CC polynucleotide sequence, where in each gene of interest are operably-
 CC linked to a pro nucleotide sequence of at least one gene of interest. The
 CC transposon-based vector comprises the polynucleotide and further
 CC comprises a transposase gene operably linked to a first promoter and
 CC where the first promoter comprises a modified Kozak sequence, two or more
 CC genes of interest are each operably-linked to one or more additional
 CC promoters, and two or more genes of interest and their operably-linked
 CC promoters are flanked by transposase insertion sequences recognized by a
 CC transposase encoded by the transposase gene. Many diseases and disease
 CC causing organisms can be targets of the vaccine of the present invention.
 CC Some of these include, anthrax, botulism, brucellosis, glanders, Q fever,
 CC plague, shigellosis, small pox, tularemia, viral encephalitis, typhus
 CC fever, viral hemorrhagic fever, etc. The preferred animal for production
 CC of protein is a bird or a mammal. The egg or milk of these animals
 CC comprises the multimeric protein encoded by the isolated polynucleotide
 CC of the invention. Hence, these proteins are endowed with varied
 CC properties and are antibacterial, anti viral, antimicrobial, anti-
 CC inflammatory, and neuroprotective in nature. The sequence presented here
 CC is the spacer sequence.

XX Sequence 20 AA;

Query Match 100.0%; Score 112; DB 8; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
 |||||
 Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 26

ADT49619

ID ADT49619 standard; peptide; 21 AA.

XX AC ADT49619;

XX 18-NOV-2004 (first entry)

XX Human IgG1 Fc region/IFNalpha2b fusion protein-related linker peptide 2.

XX immunogenic; antiinflammatory; hepatotropic; virucide;
 KW Hepatitis C virus infection; HCV infection; interferon alpha 2b;
 KW IFNalpha2b; IgG1 hinge CH2 CH3 Fc region; linker.

XX Synthetic.

XX WO2004074486-A2.

XX 02-SEP-2004.

XX 18-FEB-2004; 2004WO-EP001524.

XX 18-FEB-2003; 2003EP-00003647.

XX (MERE) MERCK PATENT GMBH.

XX Jones T, Baker M, Hanlon M, Carr FJ;

XX WPI; 2004-635581/61.

XX New modified human Interferon alpha 2 having improved biological and
 PT immunogenic properties, useful for treating hepatitis C virus infection.

AA94379
ID AAY94379 standard; peptide; 24 AA.
XX AC AAY94379;
XX DT 22-AUG-2000 (first entry)
XX DE Zinc finger stalling polypeptide sequence.
XX KW Zinc finger; DNA binding protein; ribosome stalling; polysome display;
XX KW binding protein selection.
XX OS Unidentified.
XX PN WO200027878-A1.
XX PD 18-MAY-2000.
XX PF 09-NOV-1999; 99WO-GB003730.
XX PR 09-NOV-1998; 98GB-00024544.
XX PA (GEND-) GENDAQ LTD.
XX PI Choo Y;
XX DR WPI; 2000-376494/32.
XX PT A polysome display-based technique for producing and selecting zinc
XX PT finger nucleic acid binding proteins with desired binding
XX PT characteristics.
XX PS Disclosure; Page 29; 49pp; English.
XX CC The present sequence is a stalling polypeptide sequence which forms part
XX CC of a linker/stalling sequence region that is appended to the coding
XX CC sequence of a zinc finger gene to form a zinc finger transcription unit.
XX CC In one embodiment the serine-glycine repeat unit of the present sequence
XX CC is 31 residues long and serves as a flexible linker, when translated,
XX CC which ensures that the expressed zinc finger construct has spatial
XX CC separation and flexibility with respect to the stalled ribosome. Adjacent
XX CC to this sequence there is a series of six AAVP repeats which ensure that
XX CC the ribosome is stalled after the entire flexible Ser-Gly linker has
XX CC emerged from the ribosome. By modifying the coding sequence, the zinc
XX CC finger protein may be varied at one or more positions and variants which
XX CC bind to a target nucleic acid sequence may be selected by polysome
XX CC display. In this way, zinc finger binding proteins with desired binding
XX CC characteristics can be isolated without resorting to phage display
XX CC techniques
XX SQ Sequence 24 AA;
Query Match 100.0%; Score 112; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGGSGGGSGGGSGGGSGGS 20
Db 1 GGGSGGGSGGGSGGGSGGS 20
RESULT 30
ID ADM95078 standard; peptide; 24 AA.
XX AC ADM95078;
XX DT 17-JUN-2004 (first entry)
XX DE Poly-glycine linker #2.
XX KW Cytostatic; Immunosuppressive; Antimicrobial;
XX KW major histocompatibility complex; MHC; cancer; infectious disease;
AA94379
ID AAY94379 standard; peptide; 24 AA.
XX AC AAY94379;
XX DT 22-AUG-2000 (first entry)
XX DE Zinc finger stalling polypeptide sequence.
XX KW Zinc finger; DNA binding protein; ribosome stalling; polysome display;
XX KW binding protein selection.
XX OS Unidentified.
XX PN WO200027878-A1.
XX PD 18-MAY-2000.
XX PF 09-NOV-1999; 99WO-GB003730.
XX PR 09-NOV-1998; 98GB-00024544.
XX PA (GEND-) GENDAQ LTD.
XX PI Choo Y;
XX DR WPI; 2000-376494/32.
XX PT A polysome display-based technique for producing and selecting zinc
XX PT finger nucleic acid binding proteins with desired binding
XX PT characteristics.
XX PS Disclosure; Page 29; 49pp; English.
XX CC The present sequence is a stalling polypeptide sequence which forms part
XX CC of a linker/stalling sequence region that is appended to the coding
XX CC sequence of a zinc finger gene to form a zinc finger transcription unit.
XX CC In one embodiment the serine-glycine repeat unit of the present sequence
XX CC is 31 residues long and serves as a flexible linker, when translated,
XX CC which ensures that the expressed zinc finger construct has spatial
XX CC separation and flexibility with respect to the stalled ribosome. Adjacent
XX CC to this sequence there is a series of six AAVP repeats which ensure that
XX CC the ribosome is stalled after the entire flexible Ser-Gly linker has
XX CC emerged from the ribosome. By modifying the coding sequence, the zinc
XX CC finger protein may be varied at one or more positions and variants which
XX CC bind to a target nucleic acid sequence may be selected by polysome
XX CC display. In this way, zinc finger binding proteins with desired binding
XX CC characteristics can be isolated without resorting to phage display
XX CC techniques
XX SQ Sequence 24 AA;
Query Match 100.0%; Score 112; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGGSGGGSGGGSGGGSGGS 20
Db 1 GGGSGGGSGGGSGGGSGGS 20
RESULT 31
ID AAW18555 standard; protein; 25 AA.
XX AC AAW18555;
XX DT 27-MAR-1998 (first entry)
XX DE Linker sequence 2 for use in a novel fusion protein.
XX KW Kunitz type active site inhibitor domain; fusion protein; Factor VIIa;
XX KW FVIIa; linker; tissue factor; TF; treatment; procoagulant activity;
XX KW thrombosis.
XX OS Synthetic.
XX PN WO9720939-A1.
XX PD 12-JUN-1997.
XX PF 22-NOV-1996; 96WO-US018756.
XX PR 01-DEC-1995; 95US-00566459.

PT useful as diagnostic markers or as cell-specific targeting agents.

XX Example 2; Page 45; 156pp; English.

PS This invention relates to novel vector constructs for the expression of

XX streptavidin fusion proteins. Streptavidin (SA) is produced by

CC Streptomyces avidinii and specifically binds water-soluble biotin. The

CC vectors comprise a nucleic acid sequence encoding genomic streptavidin, a

CC promoter operatively linked to the nucleic acid sequence encoding an anti-

CC site for insertion of a second nucleic acid sequence encoding an anti-

CC CD25 antibody or its antigen binding fragment. The fusion proteins

CC encoded are known as single chain antibody-genomic streptavidin fusion

CC proteins (scFvSA). The vectors may have cytostatic activity when used in

CC gene therapy. The vectors may be useful in expressing genomic

CC streptavidin fusion cassettes. The fusion proteins may be used as

CC diagnostic markers or as cell-specific targeting agents. These may also

CC be used in treating tumours. The present sequence is that of synthetic

CC amino acid linker Seq ID11 used in the streptavidin fusion proteins of

CC the invention

XX

SQ Sequence 25 AA;

Query Match 100.0%; Score 112; DB 6; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20

Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 34

ABU64228

ID ABU64228 standard; peptide; 25 AA.

AC ABU64228;

XX

DT 11-MAR-2004 (first entry)

XX

DE HIV gp 160 antigenic composition linker peptide #5.

XX

XX HIV; envelope glycoprotein; gp 160; antigen; gp 120; gp 41; vaccine;

XX anti-HIV; linker.

XX Synthetic.

OS

XX WO2003077838-A2.

PN

XX

PD 25-SEP-2003.

XX

XX 05-MAR-2002; 2002WO-US007144.

PF

XX

XX 05-MAR-2002; 2002WO-US007144.

PR

XX (NAHE-) NAT INST HEALTH OFFICE TECHNOLOGY TRANSF.

PA

XX Dimitrov DS, Chow Y, Phogat SK, Broder CC;

PI

XX WPI; 2003-779074/73.

DR

XX New HIV antigenic composition comprising an HIV envelope glycoprotein 160

PT having a glycoprotein 120 subunit covalently linked to a glycoprotein 41

PT subunit useful for protecting a human from HIV infection.

PS

XX Claim 4; Page 65; 65pp; English.

PS

XX The present invention relates to an HIV antigenic composition, comprising

CC an HIV envelope glycoprotein (gp) 160 having a gp120 subunit and a gp41

CC subunit, where the carboxy-terminal end of gp120 is covalently linked

CC through a peptide linker of at least 5 amino acids, to the amino-terminal

CC end of gp41. The HIV antigenic composition is useful for protecting a

CC human from HIV infection, potentially in the form of a vaccine. The

CC present sequence is a linker peptide used to construct the antigenic form

CC of the HIV gp160 protein sequence

XX

SQ Sequence 25 AA;

Query Match 100.0%; Score 112; DB 7; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20

Db 4 GGGGGGGGGGGGGGGGGG 23

RESULT 35

ADM95077

ID ADM95077 standard; peptide; 25 AA.

XX

AC ADM95077;

XX

DT 17-JUN-2004 (first entry)

XX

DE Poly-glycine linker #1.

XX

XX Cytostatic; Immunosuppressive; Antimicrobial;

KW major histocompatibility complex; MHC; cancer; infectious disease;

KW autoimmune disease; transplant rejection prevention.

XX

OS Synthetic.

OS

XX GB2392158-A.

PN

XX

PD 25-FEB-2004.

XX

XX 21-AUG-2002; 2002GB-00019459.

PF

XX

XX 21-AUG-2002; 2002GB-00019459.

PR

XX (PROI-) PROIMMUNE LTD.

PA

XX Schwabe NFG, Tan LC, Napper CE, Pang S, Fry JW, Spooner RK;

PI

XX WPI; 2004-193388/19.

DR

XX

XX New oligomeric major histocompatibility complex (MHC) comprising a first

PT section derived from an MHC peptide chain, and a second section

PT comprising an oligomerizing domain, useful for treating e.g. cancers or

PT autoimmune diseases.

XX

PS Example 1.3; Page 24; 48pp; English.

XX

XX The present invention relates to a novel oligomeric major

CC histocompatibility complex (MHC) comprising at least 2 chimeric proteins,

CC comprising a first section derived from an MHC peptide chain or its

CC functional part, and a second section comprising an oligomerizing domain

CC derived from an oligomer-forming coiled-coil protein. The oligomeric MHC

CC may be used therapeutically for treatment of cancers, infectious

CC diseases, autoimmune diseases, and prevention of transplant rejection; or

CC in developing and performing assay procedures or immunostainings. The

CC complexes may also be used for the separation, purification and

CC enrichment of T cells that are specific for a particular antigen.

CC Oligonucleotides ADM95069 and ADM95070 were used to synthesise a poly-

CC glycine linker (ADM95077) for use in an example from the invention.

XX

SQ Sequence 25 AA;

Query Match 100.0%; Score 112; DB 8; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20

Db 3 GGGGGGGGGGGGGGGGGG 22

RESULT 36
 ADO21852
 ID ADO21852 standard; peptide; 25 AA.
 XX
 AC ADO21852;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Histone acetylation fusion protein reporter-related linker peptide 13.
 XX
 KW fusion protein reporter; histone modification; cytostatic;
 KW neuroprotective; muscular; virucide; cancer; neurological;
 KW muscular disorders; developmental; cell development; toxin;
 KW toxic chemical; virus; linker.
 XX
 OS Synthetic.
 XX
 PN WO200404168-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 12-NOV-2003; 2003WO-US036059.
 XX
 PR 12-NOV-2002; 2002US-0425578P.
 PR 05-AUG-2003; 2003US-00634740.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Ting AY;
 XX
 DR WPI; 2004-420310/39.
 XX
 PT New fusion protein reporter having a core with a histone-modification-
 PT specific binding domain conjugated to a histone polypeptide, useful for
 PT diagnosing or treating cancer, viral infections, neurological and
 PT muscular disorders.
 XX
 PS Disclosure; SEQ ID NO 21; 96pp; English.
 XX
 PI Ting AY;
 XX
 DR WPI; 2004-420310/39.
 XX
 PT New fusion protein reporter having a core with a histone-modification-
 PT specific binding domain conjugated to a histone polypeptide, useful for
 PT diagnosing or treating cancer, viral infections, neurological and
 PT muscular disorders.
 XX
 PS Disclosure; SEQ ID NO 21; 96pp; English.
 XX
 PI The invention relates to a novel fusion protein reporter comprising a
 CC core having a histone-modification-specific binding domain conjugated to
 CC a histone polypeptide, wherein the core is flanked by donor and acceptor
 CC fluorescent moieties. The molecules of the invention demonstrate
 CC cytosolic, neuroprotective, muscular and virucide activities and may be
 CC useful for the diagnosis and/or treatment of histone modification-
 CC associated disorders, such as cancer, neurological disorders, muscular
 CC disorders, developmental disorders, cell development disorders and
 CC disorders related to exposure to toxins, toxic chemicals or viruses. The
 CC current sequence is that of the histone acetylation fusion protein
 CC reporter-related linker peptide 13 of the invention.
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 112; DB 8; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 DB 1 GGGGGGGGGGGGGGGGGG 20
 RESULT 37
 ADQ77071
 ID ADQ77071 standard; peptide; 25 AA.
 XX
 AC ADQ77071;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Peptide linker #2.
 XX

KW Antiinflammatory; Antiallergic; Cytostatic; Antibacterial; Virucide;
 KW Immunosuppressive; Immunotherapy; single domain antibody; dAb; TNF alpha;
 KW TNF receptor 1; p55; Inflammation; allergic hypersensitivity; cancer;
 KW bacterial infection; viral infection; autoimmune disorder.
 XX
 OS Synthetic.
 XX
 PN WO2004058821-A2.
 XX
 PD 15-JUL-2004.
 XX
 PF 24-DEC-2003; 2003WO-GB005646.
 XX
 PR 27-DEC-2002; 2002GB-00030202.
 PR 30-JUN-2003; 2003WO-GB002804.
 PR 28-NOV-2003; 2003GB-00027706.
 XX
 PA (DOMA-) DOMANTIS LTD.
 XX
 PI Winter G, Tomlinson I, Ignatovich O, Woolven B;
 XX
 DR WPI; 2004-534127/51.
 XX
 PT New dual specific ligand comprising a first dAb specific for a target
 PT ligand, and a second dAb specific for a receptor for the target ligand,
 PT useful for treating inflammation, cancer, allergy or autoimmune diseases.
 XX
 PS Example 6; Page 98; 196pp; English.
 XX
 PI The present invention relates to a dual specific ligand comprising a
 CC first single domain antibody (dAb) specific for a target ligand (e.g. TNF
 CC alpha), and a second dAb specific for a receptor for the target ligand
 CC (e.g. TNF receptor 1 (p55)). The dAb specific for TNFalpha comprises the
 CC amino acid sequence of TAR1-5-19, TAR1-5 or TAR1-27 and the dAb specific
 CC for TNF receptor 1 (p55) comprises the amino acid sequence of TAR2h-10,
 CC TAR2h-5 or TAR2h-10-27. The dual specific ligands are useful for
 CC targeting cytokines and other molecules that cooperate synergistically in
 CC therapeutic situations in the body of an organism. They are useful for
 CC preventing, suppressing or treating inflammatory states, allergic
 CC disorders, hypersensitivity, cancer, bacterial or viral infection, and autoimmune
 CC disorders. The present sequence is a peptide linker used in an example
 XX from the invention.
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 112; DB 8; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGGGGGGG 20
 DB 1 GGGGGGGGGGGGGGGGGG 20
 RESULT 38
 ADL24301
 ID ADL24301 standard; protein; 26 AA.
 XX
 AC ADL24301;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Cardiovascular disease treatment related linker #5.
 XX
 KW cardiovascular disease; IL-17; interleukin; IL-18; 4-1BB; CD30; OX40;
 KW antagonist.
 XX
 OS Synthetic.
 XX
 PN WO2004019866-A2.
 XX
 PD 11-MAR-2004.
 XX

PF 21-AUG-2003; 2003WO-US026354.
XX
PR 28-AUG-2002; 2002US-0406418P.
PR 12-AUG-2003; 2003US-0494457P.
XX
XX (IMMUNEX CORP.
XX
XX Burton PB, Deisher TA;
XX
XX WPI; 2004-239107/22.
DR
XX
XX Use of IL-17, IL-18, 4-1BB, CD30 or OX40 antagonists, for treating a
XX cardiovascular disease, e.g. chronic immune myocarditis, congestive heart
PT failure, aneurysm, angina, embolism, restenosis, ischemia or
PT thrombocytopenic purpura.
XX
XX Disclosure; Page 135; 135pp; English.
PS
XX The present invention relates to a method of treating cardiovascular
CC disease in a subject, comprising administering an IL-17, IL-18, 4-1BB,
CC CD30 or OX40 antagonist. The IL-17, IL-18, 4-1BB, CD30 or OX40
CC antagonists are useful for treating cardiovascular disorders, e.g.
CC (chronic immune) myocarditis, congestive heart failure, aneurysms,
CC angina, embolism, restenosis, ischaemia or thrombocytopenic purpura. The
CC present sequence is a polypeptide used in the exemplification of the
CC invention.
XX
XX Sequence 26 AA;
SQ
Query Match 100.0%; Score 112; DB 8; Length 26;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGSGGGGGGGGGGGGGG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGGGGGGGGGGGG 20
| | | | | | | | | | | | | | | | | | | | | |
RESULT 39
AAY25360
ID AAY25360 standard; peptide; 28 AA.
XX
XX AAY25360;
AC
XX 06-SEP-1999 (first entry)
DT
XX IFNAR2/IFN-beta complex peptide fragment 3.
DE
XX IFNAR2; IFN-beta; type I interferon; IFNAR/IFN complex; IFN; antiviral;
KW human interferon alpha/beta receptor; anticancer; immunomodulatory;
KW anti-arthritis; antidiabetic; treatment; hepatitis; viral infection;
KW hairy cell leukemia; Kaposi's sarcoma; multiple myeloma; cancer; lupus;
KW diabetes; multiple sclerosis; rheumatoid arthritis; myasthenia gravis;
KW acquired immune deficiency syndrome.
XX
XX Synthetic.
OS
XX WO9932141-A1.
FN
XX 01-JUL-1999.
PD
XX 18-DEC-1998; 98WO-US026926.
PF
XX 19-DEC-1997; 97US-0068295P.
PR
XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN/) MCINIS P G.
PA
XX Pepper M, Cunningham M, Sherris D, El Tayar N, McKenna S;
PI WPI; 1999-405115/34.
XX
XX Prolonging in vivo activity of type I interferon by complexing.
PT

XX
PS Example 8; Page 77; 86pp; English.
XX
CC This invention describes a novel method for prolonging the in vivo effect
CC of type I interferon (IFN) by administering IFN as a complex (A) with a
CC subunit (I) of the human interferon alpha/beta receptor (IFNAR). The
CC product of the invention has antiviral, anticancer, immunomodulatory,
CC anti-arthritis and antidiabetic activity. (A) have the antiviral,
CC anticancer and immunomodulating activities of IFN, e.g. for treating
CC hepatitis and other viral infections, hairy cell leukemia, Kaposi's
CC sarcoma, multiple myeloma and other cancers, multiple sclerosis,
CC rheumatoid arthritis, myasthenia gravis, diabetes, acquired immune
CC deficiency syndrome and lupus. When complexed in (A), the storage life of
CC IFN is increased (i.e. it is stabilized against oligomerization, without
CC the need for storage at acidic pH) and its biological effect is
CC potentiated
XX
XX Sequence 28 AA;
SQ
Query Match 100.0%; Score 112; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGSGGGGGGGGGGGGGG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 6 GGGSGGGGGGGGGGGGGG 25
| | | | | | | | | | | | | | | | | | | | | |
RESULT 40
AAG77868
ID AAG77868 standard; protein; 30 AA.
XX
XX AAG77868;
AC
XX 08-MAY-2002 (first entry)
DT
XX MHC class II H domain spacer region.
DE
XX MHC class II H domain spacer; IGG3-HLA fusion protein;
KW major histocompatibility complex; MHC-peptide complex; MHC;
KW MHC class I alpha chain; beta-2 microglobulin; MHC class II alpha chain;
KW MHC class II beta chain; vaccine; immune response modulation;
KW hyperproliferative disorder; neoplasm; hypermaglobulinaemia;
KW viral infection; hepatitis; meningitis; bacterial infection;
KW tuberculosis; gingivitis; parasitic infection; autoimmune disease;
KW Hashimoto's disease; Graves' disease; rheumatoid arthritis; allergy;
KW asthma; organ rejection; graft-versus-host disease; GVHD; HLA;
KW human leukocyte antigen.
XX
XX Synthetic.
OS
XX WO200178768-A2.
FN
XX 25-OCT-2001.
PD
XX 12-APR-2001; 2001WO-US011912.
PF
XX 12-APR-2000; 2000US-0196472P.
PR
XX (UYRP) UNIV ROCHESTER.
PA
XX Zauderer M, Smith ES;
PI WPI; 2001-602927/68.
DR
XX Novel compound comprising major histocompatibility complex-peptide
PT complexes, used to modulate immune responses.
PT
XX Example 3; Page 146; 166pp; English.
PS
XX The present sequence represents a major histocompatibility complex (MHC)
CC class II H domain spacer region. This protein represents a region of the
CC protein used as a replacement for the H domain present in MHC class II
CC

CC molecules. The H domain replacement protein was used in an example of the
CC invention detailing the construction of human IgG3-HLA fusion proteins.
CC Absence of the H domain or a suitable spacer would result in misfolding
CC of the MHC class II molecule and the correct MHC binding site would not
CC be formed. The invention comprises a compound which contains one or more
CC major histocompatibility complex (MHC)-peptide complexes, and an antibody
CC specific for a cell surface marker. The complexes comprise an MHC class I
CC alpha chain, a beta-2 microglobulin molecule and an antigenic peptide
CC bound in the MHC groove. Alternatively, the complexes may comprise an MHC
CC class II alpha chain, an MHC class II beta chain, and an antigenic
CC peptide bound in the MHC groove. The complexes are linked to the carboxyl
CC terminus of the antibody. The compounds of the invention can be used as a
CC vaccine to modulate an immune response. The compounds of the invention
CC are useful for treating: hyperproliferative disorders (e.g. neoplasms and
CC hypergammaglobulinaemia); viral infections (e.g. hepatitis and
CC meningitis); bacterial infections (e.g. tuberculosis and gingivitis);
CC parasitic infections; autoimmune diseases (e.g. Hashimoto's disease,
CC Graves' disease and rheumatoid arthritis); allergic reactions/conditions
CC (e.g. asthma). The compounds of the invention may also be used in the
CC treatment of organ rejection or graft-versus-host disease (GVHD)
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 112; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGSGGGSGGGSGGGSGGGGS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GGGSGGGSGGGSGGGSGGGGS 20

Search completed: August 19, 2005, 16:32:27
Job time : 165 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 16:26:59 ; Search time 41 Seconds

(without alignments)
36.414 Million cell updates/sec

Title: US-09-674-794A-14

Perfect score: 112

Sequence: 1 GGGGGGGGGGGGGGGG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA:*

1: /cgm2_6/ptodata/1/iaa/5A COMB.pcp.*

2: /cgm2_6/ptodata/1/iaa/5B COMB.pcp.*

3: /cgm2_6/ptodata/1/iaa/6A COMB.pcp.*

4: /cgm2_6/ptodata/1/iaa/6B COMB.pcp.*

5: /cgm2_6/ptodata/1/iaa/PCTUS COMB.pcp.*

6: /cgm2_6/ptodata/1/iaa/backfilea1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	20	1	US-08-225-224-55
2	112	100.0	20	3	US-08-722-258-55
3	112	100.0	20	3	US-09-046-985-14
4	112	100.0	20	3	US-09-474-743-14
5	112	100.0	20	3	US-09-178-115-116
6	112	100.0	20	3	US-09-177-776-116
7	112	100.0	20	3	US-09-425-585-12
8	112	100.0	20	4	US-09-603-663-75
9	112	100.0	20	4	US-09-603-658-75
10	112	100.0	20	4	US-09-603-372A-75
11	112	100.0	20	4	US-09-422-375-1
12	112	100.0	20	4	US-09-953-321-12
13	112	100.0	20	5	PCT-US95-04468-55
14	112	100.0	22	1	US-08-225-224-56
15	112	100.0	22	3	US-08-722-258-56
16	112	100.0	22	5	PCT-US95-04468-56
17	112	100.0	24	2	US-08-902-623-16
18	112	100.0	24	4	US-09-851-271A-12
19	112	100.0	25	4	US-09-254-832B-22
20	112	100.0	28	3	US-09-215-212-4
21	112	100.0	33	3	US-09-215-212-3
22	112	100.0	40	3	US-08-815-190A-17
23	112	100.0	40	4	US-09-333-213-1
24	112	100.0	60	4	US-09-832-297A-12
25	112	100.0	100	4	US-09-411-067C-4
26	112	100.0	355	5	PCT-US95-03866-32
27	112	100.0	456	4	US-09-495-880A-11

112	100.0	482	4	US-09-509-031-16	Sequence 16, Appl
112	100.0	500	3	US-08-960-190A-25	Sequence 25, Appl
108	96.4	20	3	US-08-851-749-6	Sequence 6, Appl
108	96.4	20	4	US-09-333-213-3	Sequence 3, Appl
108	96.4	20	4	US-10-077-210-3	Sequence 3, Appl
106	94.6	20	3	US-09-025-769B-1	Sequence 1, Appl
106	94.6	20	4	US-09-490-070A-1	Sequence 1, Appl
106	94.6	20	4	US-09-490-153-1	Sequence 1, Appl
106	94.6	20	4	US-09-490-324-1	Sequence 1, Appl
106	94.6	21	3	US-08-851-749-1	Sequence 1, Appl
106	94.6	281	3	US-09-025-769B-178	Sequence 178, App
106	94.6	281	4	US-09-490-070A-178	Sequence 178, App
106	94.6	281	4	US-09-490-153-178	Sequence 178, App
106	94.6	281	4	US-09-490-324-178	Sequence 178, App
104	92.9	979	4	US-09-538-092-990	Sequence 990, App
104	92.9	3730	4	US-09-949-016-990B	Sequence 990B, Ap
103	92.0	738	3	US-08-864-038A-3	Sequence 3, Appl
98	87.5	22	4	US-08-406-824A-24	Sequence 24, Appl
97	86.6	161	2	US-08-581-528A-6	Sequence 6, Appl
97	86.6	161	4	US-09-412-791D-6	Sequence 6, Appl
97	86.6	161	5	PCT-US94-07799-6	Sequence 6, Appl
97	86.6	240	1	US-08-362-670B-30	Sequence 30, Appl
97	86.6	240	3	US-08-333-576C-30	Sequence 30, Appl
97	86.6	240	3	US-08-808-324-30	Sequence 30, Appl
97	86.6	240	4	US-09-945-182-30	Sequence 30, Appl
97	86.6	240	5	PCT-US94-14030A-30	Sequence 30, Appl
97	86.6	249	4	US-09-370-767-44857	Sequence 44857, A
97	86.6	268	2	US-08-835-099A-9	Sequence 9, Appl
97	86.6	268	3	US-09-157-349-9	Sequence 9, Appl
96	85.7	20	1	US-08-234-602-14	Sequence 14, Appl
96	85.7	20	3	US-09-039-780A-7	Sequence 7, Appl
96	85.7	33	3	US-09-039-780A-24	Sequence 24, Appl
96	85.7	68	4	US-09-107-433-4787	Sequence 4787, Ap
96	85.7	100	4	US-09-893-600-3	Sequence 3, Appl
96	85.7	200	4	US-08-989-981A-13	Sequence 13, Appl
96	85.7	201	3	US-09-052-995-1	Sequence 1, Appl
96	85.7	201	3	US-09-053-003-40	Sequence 40, Appl
96	85.7	201	4	US-09-054-281-22	Sequence 22, Appl
96	85.7	201	4	US-09-478-948-6	Sequence 6, Appl
96	85.7	201	4	US-09-818-094-40	Sequence 40, Appl
96	85.7	201	4	US-09-754-947-5	Sequence 5, Appl
96	85.7	232	4	US-09-555-352-10	Sequence 10, Appl
96	85.7	342	3	US-08-828-741B-6	Sequence 6, Appl
96	85.7	342	3	US-09-160-567-6	Sequence 6, Appl
96	85.7	342	4	US-09-710-299-6	Sequence 6, Appl
96	85.7	342	4	US-09-509-031-6	Sequence 6, Appl
96	85.7	443	4	US-09-538-092-966	Sequence 966, App
96	85.7	443	4	US-09-949-016-6641	Sequence 6641, Ap
96	85.7	493	4	US-09-949-016-11429	Sequence 11429, A
96	85.7	493	3	US-08-828-741B-4	Sequence 4, Appl
96	85.7	495	3	US-09-160-567-4	Sequence 4, Appl
96	85.7	495	4	US-09-710-299-4	Sequence 4, Appl
96	85.7	495	4	US-09-509-031-4	Sequence 11, Appl
96	85.7	918	3	US-09-041-886-11	Sequence 895, App
96	85.7	919	4	US-09-538-092-895	Sequence 19, Appl
96	85.7	923	4	US-09-497-822C-19	Sequence 9, Appl
96	85.7	1084	4	US-09-394-272-9	Sequence 6769, Ap
95	84.8	331	4	US-09-949-016-6769	Sequence 6769, Ap
95	84.8	332	1	US-08-014-943A-2	Sequence 2, Appl
94	83.9	322	1	US-08-486-421-3	Sequence 3, Appl
94	83.9	322	1	US-08-470-911-3	Sequence 3, Appl
94	83.9	322	2	US-08-486-809-3	Sequence 3, Appl
94	83.9	345	4	US-09-657-013-112	Sequence 112, App
94	83.9	485	4	US-09-949-016-6557	Sequence 6557, Ap
94	83.9	504	4	US-09-949-016-6783	Sequence 7783, Ap
93.5	83.5	647	2	US-08-770-761A-8	Sequence 8, Appl
93.5	83.5	705	2	US-08-770-761A-7	Sequence 7, Appl
92.5	82.6	272	4	US-08-756-416-39	Sequence 39, Appl
92	82.1	21	2	US-08-484-905-14	Sequence 14, Appl
92	82.1	21	3	US-08-481-985B-14	Sequence 14, Appl
92	82.1	260	4	US-08-370-476-14	Sequence 14, Appl
91	81.2	246	1	US-09-270-767-43301	Sequence 43301, A
				US-07-843-125-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-225-224-55
; Sequence 55, Application US/08225224
; Patent No. 5635599
; GENERAL INFORMATION:
; APPLICANT: PASTAN, Ira
; APPLICANT: KREITMAN, Robert J.
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
; TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,224
; FILING DATE: 8-APR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-225-224-55
Query Match 100.0%; Score 112; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGG 20
RESULT 2
US-08-722-258-55
; Sequence 55, Application US/08722258
; Patent No. 6011002
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Kreitman, Robert J.
; APPLICANT: Puri, Raj K.
; TITLE OF INVENTION: Circularly Permutated Ligands and
; TITLE OF INVENTION: Circularly Permutated Chimeric Molecules
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA

ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,258
; FILING DATE: 08-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US95/04468
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-193100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-722-258-55
Query Match 100.0%; Score 112; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGG 20
RESULT 3
US-09-046-985-14
; Sequence 14, Application US/09046985
; Patent No. 6121236
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: NOVEL MULTIVALENT LIGANDS WHICH MODULATE
; TITLE OF INVENTION: ANGIOGENESIS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,985
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CMCC-614
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540


```
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 20 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-09-046-985-14

Query Match          100.0%; Score 112; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 4
US-09-474-743-14
; Sequence 14, Application US/09474743
; Patent No. 6235716
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: NOVEL MULTIVALENT LIGANDS WHICH MODULATE
;   TITLE OF INVENTION: ANGIOGENESIS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/474,743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/046,985
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CMCC-614
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 20 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-09-474-743-14

Query Match          100.0%; Score 112; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 5
US-09-178-115-116
```

```
; Sequence 116, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/178,115
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 09/177,776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 20
; TYPE: PRT
; ORGANISM: HUMAN
US-09-178-115-116

Query Match          100.0%; Score 112; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 6
US-09-177-776-116
; Sequence 116, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
```

```
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 20
; TYPE: PRT
; ORGANISM: HUMAN
; OTHER INFORMATION: Description of Artificial Sequence: Peptide Linker
US-09-177-776-116

Query Match      100.0%; Score 112; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 7
US-09-425-585-12
; Sequence 12, Application US/09425585
; Patent No. 6348315
; GENERAL INFORMATION:
; APPLICANT: PLUCKTHUN, ANDREAS
; APPLICANT: HANES, JOZEF
; TITLE OF INVENTION: METHODS FOR IDENTIFYING NUCLEIC ACID MOLECULES ENCODING
; FILE REFERENCE: (POLY) PEPTIDES THAT INTERACT WITH TARGET MOLECULES
; CURRENT APPLICATION NUMBER: US/09/425,585
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: PCT/EP98/02420
; PRIOR FILING DATE: 1998-04-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide Linker
US-09-425-585-12

Query Match      100.0%; Score 112; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 8
US-09-603-663-75
; Sequence 75, Application US/09603663
; Patent No. 6406863
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing
; TITLE OF INVENTION: GENERATION OF HIGHLY DIVERSE LIBRARY OF EXPRESSION
; FILE REFERENCE: VECTORS VIA HOMOLOGOUS RECOMBINATION IN YEAST
; CURRENT APPLICATION NUMBER: US/09/602,373A
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Linker peptide
US-09-603-663-75

Query Match      100.0%; Score 112; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 9
US-09-603-658-75
; Sequence 75, Application US/09603658
; Patent No. 6410246
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing
; TITLE OF INVENTION: HIGHLY DIVERSE LIBRARY OF YEAST EXPRESSION VECTORS
; FILE REFERENCE: 25636-703 Seq Listing
; CURRENT APPLICATION NUMBER: US/09/603,658
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Linker peptide
US-09-603-658-75

Query Match      100.0%; Score 112; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 10
US-09-602-373A-75
; Sequence 75, Application US/09602373A
; Patent No. 6410271
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing B.
; TITLE OF INVENTION: GENERATION OF HIGHLY DIVERSE LIBRARY OF EXPRESSION
; FILE REFERENCE: VECTORS VIA HOMOLOGOUS RECOMBINATION IN YEAST
; CURRENT APPLICATION NUMBER: US/09/602,373A
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Linker peptide
US-09-602-373A-75

Query Match      100.0%; Score 112; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20
```

; OTHER INFORMATION: Description of Artificial Sequence: Linker peptide
US-09-602-373A-75

Query Match 100.0%; Score 112; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 11

US-09-422-375-1
; Sequence 1, Application US/09422375
; Patent No. 6534633
; GENERAL INFORMATION:
; APPLICANT: Weidanz, Jon A.
; APPLICANT: Card, Kimberlyn F.
; APPLICANT: Sherman, Linda A.
; APPLICANT: Klinman, No. 6534633man
; APPLICANT: Wong, Hing C.
; TITLE OF INVENTION: POLYSPECIFIC BINDING MOLECULES AND USES THEREOF
; FILE REFERENCE: 48531
; CURRENT APPLICATION NUMBER: US/09/422.375
; CURRENT FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-422-375-1

Query Match 100.0%; Score 112; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 12

US-09-953-321-12
; Sequence 12, Application US/09953321
; Patent No. 6589741
; GENERAL INFORMATION:
; APPLICANT: PLUCKTHUN, ANDREAS
; APPLICANT: HANES, JOZEF
; APPLICANT: JERMUTUS, LUTZ
; TITLE OF INVENTION: NOVEL METHODS FOR IDENTIFYING NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: (POLY) PEPTIDES THAT INTERACT WITH TARGET MOLECULES
; FILE REFERENCE: PLUCK/1 CON2
; CURRENT APPLICATION NUMBER: US/09/953.321
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 09/425,585
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide Linker
US-09-953-321-12

Query Match 100.0%; Score 112; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20

Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 13

PCT-US95-04468-55
; Sequence 55, Application PC/TUS9504468
; GENERAL INFORMATION:
; APPLICANT: CIRCULARLY PERMUTATED LIGANDS AND
; TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS
; NUMBER OF SEQUENCES: 59
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04468
; FILING DATE: 07-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/225,224
; APPLICATION NUMBER: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-193-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-04468-55

Query Match 100.0%; Score 112; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGSGGGSGGGGS 20
Db 1 GGGSGGGSGGGSGGGGS 20

RESULT 14

US-08-225-224-56
; Sequence 56, Application US/08225224
; Patent No. 5635599
; GENERAL INFORMATION:
; APPLICANT: PASTAN, Ira
; APPLICANT: KREITMAN, Robert J.
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
; TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,224

```
; FILING DATE: 8-APR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-225-224-56

Query Match 100.0%; Score 112; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
Db 3 GGGSGGGSGGGSGGGGS 22

RESULT 15
US-08-722-258-56
; Sequence 56, Application US/08722258
; Patent No. 6011002
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Kreitman, Robert J.
; APPLICANT: Puri, Raj K.
; TITLE OF INVENTION: Circularly Permuted Ligands and
; TITLE OF INVENTION: Circularly Permuted Chimeric Molecules
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,258
; FILING DATE: 08-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US95/04468
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-193100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: peptide
US-08-722-258-56

Query Match 100.0%; Score 112; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
Db 3 GGGSGGGSGGGSGGGGS 22

RESULT 16
PCT-US95-04468-56
; Sequence 56, Application PC/TUS9504468
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
; TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS
; NUMBER OF SEQUENCES: 59
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04468
; FILING DATE: 07-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-193-1PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-04468-56

Query Match 100.0%; Score 112; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGSGGGSGGGGS 20
Db 3 GGGSGGGSGGGSGGGGS 22

RESULT 17
US-08-902-623-16
; Sequence 16, Application US/08902623
; Patent No. 5922545
; GENERAL INFORMATION:
; APPLICANT: MATHEAKIS, LARRY C.
; APPLICANT: DOWER, WILLIAM J.
; TITLE OF INVENTION: IN VITRO PEPTIDE AND ANTIBODY DISPLAY
; TITLE OF INVENTION: LIBRARIES
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
```

```

; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902.623
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/586.176
; FILING DATE: 17-JAN-1996
; APPLICATION NUMBER: US 08/300.262
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/144.775
; FILING DATE: US 29-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12206
; FILING DATE: US 25-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DUNN, TRACY J.
; REGISTRATION NUMBER: 34,587
; REFERENCE/DOCKET NUMBER: 16528X-003230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-902-623-16

Query Match 100.0%; Score 112; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 18
US-09-851-271A-12
; Sequence 12, Application US/09851271A
; Patent No. 6733970
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851.271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(24)
; OTHER INFORMATION: linker sequence followed by the stalling polypeptide sequence
US-09-851-271A-12

Query Match 100.0%; Score 112; DB 4; Length 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 19
US-09-254-832B-22
; Sequence 22, Application US/09254832B
; Patent No. 6541219
; GENERAL INFORMATION:
; APPLICANT: KINGSMAN, ALAN J.
; APPLICANT: KINGSMAN, SUSAN M.
; TITLE OF INVENTION: THERAPEUTIC GENE
; FILE REFERENCE: 9192.10USWO
; CURRENT APPLICATION NUMBER: US/09/254.832B
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/GB97/02969
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 9622500.8
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-254-832B-22

Query Match 100.0%; Score 112; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGG 20
Db 1 GGGGGGGGGGGGGGGGGG 20

RESULT 20
US-09-215-212-4
; Sequence 4, Application US/09215212
; Patent No. 6372207
; GENERAL INFORMATION:
; APPLICANT: TEPPER, Mark
; APPLICANT: CUNNINGHAM, Mark
; APPLICANT: SHERRIS, David
; APPLICANT: EL TAYAR, Nabil
; APPLICANT: MCKENNA, Sean
; TITLE OF INVENTION: IFNAR2/IFN COMPLEX
; FILE REFERENCE: TEPPERIA.SEQ
; CURRENT APPLICATION NUMBER: US/09/215.212
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068.295
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C terminal human IFNbeta
US-09-215-212-4

Query Match 100.0%; Score 112; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GGGGGGGGGGGGGGGGGGS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 6 GGGGGGGGGGGGGGGGGGS 25

RESULT 21
US-09-215-212-3
; Sequence 3, Application US/09215212
; Patent No. 6372207
; GENERAL INFORMATION:
; APPLICANT: TEPPER, Mark
; APPLICANT: CUNNINGHAM, Mark
; APPLICANT: SHERRIS, David
; APPLICANT: EL TAYAR, Nabil
; APPLICANT: MCKENNA, Sean
; TITLE OF INVENTION: IFNAR2/IFN COMPLEX
; FILE REFERENCE: TEPPERIA.SQ
; CURRENT APPLICATION NUMBER: US/09/215,212
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/068,295
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C terminal human
; OTHER INFORMATION: sIFNAR2 linked by linker to N terminal human IFNbeta
US-09-215-212-3

Query Match 100.0%; Score 112; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 6 GGGGGGGGGGGGGGGGGGS 25

RESULT 22
US-08-815-190A-17
; Sequence 17, Application US/08815190A
; Patent No. 6046310
; GENERAL INFORMATION:
; APPLICANT: Queen, Cary L.
; APPLICANT: Schneider, William P.
; APPLICANT: Vasquez, Maximiliano
; TITLE OF INVENTION: Fas Ligand Fusion Proteins and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/815,190A
; FILING DATE: 11-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/614,584
; FILING DATE: 13-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.

; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 011823-006710US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 6..40
; OTHER INFORMATION: /note="amino acid residues 6-40 may be
; OTHER INFORMATION: present or absent"
US-08-815-190A-17

Query Match 100.0%; Score 112; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GGGGGGGGGGGGGGGGGGS 20

RESULT 23
US-09-333-213-1
; Sequence 1, Application US/09333213
; Patent No. 6548653
; GENERAL INFORMATION:
; APPLICANT: Young, Michael
; APPLICANT: Meade, Harry
; APPLICANT: Krane, Ian
; TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
; FILE REFERENCE: 10275/041001
; CURRENT APPLICATION NUMBER: US/09/333,213
; CURRENT FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated linker sequence; subsets 2 through 8 (each
; OTHER INFORMATION: consisting of a repetition of the first five amino acids) encompass
; OTHER INFORMATION: positions 6 through 40 may be absent or present
US-09-333-213-1

Query Match 100.0%; Score 112; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 2 GGGGGGGGGGGGGGGGGGS 21

RESULT 24
US-09-832-297A-12
; Sequence 12, Application US/09832297A
; Patent No. 6652836
; GENERAL INFORMATION:
; APPLICANT: Fluoroprobe, Inc.
; APPLICANT: LUKEN, George A.
; TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY
; FILE REFERENCE: FLUOR1120-2
; CURRENT APPLICATION NUMBER: US/09/832,297A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/362,805

```

; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/173,190
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Peptide linker moiety
; NAME/KEY: REPEAT
; LOCATION: (3)..(6)
; OTHER INFORMATION: Amino Acid at residue 3 could be repeated up to 3 times
; FEATURE:
; NAME/KEY: REPEAT
; LOCATION: (3)..(57)
; OTHER INFORMATION: Amino Acids at residues 3 to 7 could be repeated up to 10 times
; US-09-832-297A-12

Query Match      100.0%; Score 112; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.3e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGS 20
   |||||
Db 3 GGGGGGGGGGGGGGGGGGGS 22

RESULT 25
US-09-411-067C-4
; Sequence 4, Application US/09411067C
; Patent No. 6576610
; GENERAL INFORMATION:
; APPLICANT: NUVAS, LLC
; TITLE OF INVENTION: USE OF A CONTEXT-DEPENDENT FUNCTIONAL ENTITY TO ENHANCE THE EFFIC
; FILE REFERENCE: NUVAS1140
; CURRENT APPLICATION NUMBER: US/09/411.067C
; CURRENT FILING DATE: 2002-05-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Spacer element
; NAME/KEY: REPEAT
; LOCATION: (1)..(5)
; OTHER INFORMATION: Spacer peptide variation; sequence repeated "n" times, where n =
; OTHER INFORMATION: 1-20
; US-09-411-067C-4

Query Match      100.0%; Score 112; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGS 20
   |||||
Db 1 GGGGGGGGGGGGGGGGGGGS 20

RESULT 26
PCT-US95-03866-32
; Sequence 32, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: CytoMed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
;

; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-03866-32

Query Match      100.0%; Score 112; DB 5; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGS 20
   |||||
Db 169 GGGGGGGGGGGGGGGGGGGS 188

RESULT 27
US-09-495-880A-11
; Sequence 11, Application US/09495880A
; Patent No. 6667150
; GENERAL INFORMATION:
; APPLICANT: RUDERT, FRITZ
; APPLICANT: GE, LIMING
; APPLICANT: ILAG, VIC
; TITLE OF INVENTION: NOVEL METHOD AND PHASE FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING MEMBERS OF A MULTIMERIC
; TITLE OF INVENTION: (POLY)PEPTIDE COMPLEX
; FILE REFERENCE: MORPHO/9
; CURRENT APPLICATION NUMBER: US/09/495,880A
; CURRENT FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: PCT/EP98/04836
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: EP 97 11 3319.4
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ompA-FLAG-scFv (anti-
; OTHER INFORMATION: HAG)-gene IIIs encoded by phage vector thagIA (circular)

```

US-09-495-880A-11

Query Match 100.0%; Score 112; DB 4; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 140 GGGGGGGGGGGGGGGGGG 159

RESULT 28

US-09-509-031-16
; Sequence 16, Application US/09509031
; Patent No. 6590080
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/09/509,031
; CURRENT FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ccMTLgI protein
; OTHER INFORMATION: sequence

US-09-509-031-16
Query Match 100.0%; Score 112; DB 4; Length 482;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 154 GGGGGGGGGGGGGGGGGG 173

RESULT 29

US-08-960-190A-25
; Sequence 25, Application US/08960190A
; Patent No. 6232445
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Acevedo, Jorge
; APPLICANT: Burkhardt, Martin
; APPLICANT: Jiao, Jin-an
; APPLICANT: Wong, Hing C.
; TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: usa
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,190A
; FILING DATE: 29-OCT-1997
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corless, Peter F
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 48002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-960-190A-25

Query Match 100.0%; Score 112; DB 3; Length 500;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 20
Db 246 GGGGGGGGGGGGGGGGGG 265

RESULT 30

US-08-851-749-6
; Sequence 6, Application US/08851749
; Patent No. 6001329
; GENERAL INFORMATION:
; APPLICANT: Buchsbaum, Donald J.
; APPLICANT: Vallera, Daniel A.
; APPLICANT: Blazer, Bruce R.
; TITLE OF INVENTION: Radiolabeled Fusion Toxins for Cancer Therapy
; FILE REFERENCE: D5854
; CURRENT APPLICATION NUMBER: US/08/851,749
; CURRENT FILING DATE: 1997-05-06
; EARLIER APPLICATION NUMBER: US 60/016,982
; EARLIER FILING DATE: 1996-05-06
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by DNA linking mGM-CSF
; OTHER INFORMATION: to DT.
; US-08-851-749-6

Query Match 96.4%; Score 108; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGG 19
Db 2 GGGGGGGGGGGGGGGGGG 20

RESULT 31

US-09-333-213-3
; Sequence 3, Application US/09333213
; Patent No. 6548653
; GENERAL INFORMATION:
; APPLICANT: Young, Michael
; APPLICANT: Meade, Harry
; APPLICANT: Krane, Ian
; TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
; FILE REFERENCE: 10275/041001


```
; CURRENT APPLICATION NUMBER: US/09/333,213
; CURRENT FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated linker sequence
US-09-333-213-3

Query Match          96.4%; Score 108; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
Db 2 GGGSGGGGGGGGGGGGGG 20

RESULT 32
US-10-077-210-3
; Sequence 3, Application US/10077210
; Patent No. 6800462
; GENERAL INFORMATION:
; APPLICANT: Wu, Chung-Hsiun
; APPLICANT: Lin, Rong-Hwa
; APPLICANT: Hsu, Pei-Ling
; TITLE OF INVENTION: PRODUCTION OF RECOMBINANT PROTEINS IN VIVO AND USE FOR GENERATING
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 13062-004001
; CURRENT APPLICATION NUMBER: US/10/077,210
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/318,474
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker sequence
US-10-077-210-3

Query Match          96.4%; Score 108; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGG 19
Db 2 GGGSGGGGGGGGGGGGGG 20

RESULT 33
US-09-025-769B-1
; Sequence 1, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESS: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
```

```
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-1

Query Match          94.6%; Score 106; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGSGGGGGGGGGGGGGG 20
Db 2 GGGSGGGGGGGGGGGGGG 20

RESULT 34
US-09-490-070A-1
; Sequence 1, Application US/09490070A
; Patent No. 6696248
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESS: Colin G. Sandercock, Esq. c/o Heller Ehrman
; ADDRESS: White & McAuliffe
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,070A
; FILING DATE: 24-Jan-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Colin G. Sandercock, Esq.
; REGISTRATION NUMBER: 31,298
```

; REFERENCE/DOCKET NUMBER: 37629-0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 912-2000
; TELEFAX: (202) 912-2020
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-490-070A-1

Query Match 94.6%; Score 106; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGSGGGSGGGSGGGGS 20
Db 2 GGGSGGGSGGGSGGGGS 20

RESULT 35
US-09-490-153-1
; Sequence 1, Application US/09490153
; Patent No. 6706484
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; Pack, Peter
; Ilag, Vic
; Ge, Liming
; Moroney, Simon
; Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,153
; FILING DATE: 24-Jan-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-490-153-1

Query Match 94.6%; Score 106; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GGGSGGGSGGGSGGGGS 20
Db 2 GGGSGGGSGGGSGGGGS 20
RESULT 36
US-09-490-324-1
; Sequence 1, Application US/09490324
; Patent No. 6828422
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; Pack, Peter
; Ilag, Vic
; Ge, Liming
; Moroney, Simon
; Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,324
; FILING DATE: 24-Jan-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769
; FILING DATE: 18-FEB-1998
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-490-324-1

Query Match 94.6%; Score 106; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGSGGGSGGGSGGGGS 20
Db 2 GGGSGGGSGGGSGGGGS 20

RESULT 37
US-08-851-749-1
; Sequence 1, Application US/08851749
; Patent No. 6001329
; GENERAL INFORMATION:

APPLICANT: Buchsbaum, Donald J.
APPLICANT: Vallera, Daniel A.
APPLICANT: Blazer, Bruce R.
TITLE OF INVENTION: Radiolabeled Fusion Toxins for Cancer Therapy
FILE REFERENCE: D5854
CURRENT APPLICATION NUMBER: US/08/851,749
CURRENT FILING DATE: 1997-05-06
EARLIER APPLICATION NUMBER: US 60/016,982
EARLIER FILING DATE: 1996-05-06
NUMBER OF SEQ ID NOS: 10
SEQ ID NO 1
LENGTH: 21
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Amino acid sequence encoded by DNA linking MGM-CSF
OTHER INFORMATION: to DT after restriction endonuclease digestion of
OTHER INFORMATION: pDT-GM-CSF.
US-08-851-749-1

Query Match 94.6%; Score 106; DB 3; Length 21;
Best Local Similarity 95.0%; Pred. No. 8.4e-05;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGSGGGGGGGGGGGGGGS 20
|||||
Db 1 GGGSGGGGGGGGGGGGGGS 20

RESULT 38
US-09-025-769B-178
Sequence 178, Application US/09025769B
Patent No. 630064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
APPLICANT: Plueckthun, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-025-769B-178

Query Match 94.6%; Score 106; DB 3; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGSGGGGGGGGGGGGS 20
|||||
Db 147 GGGSGGGGGGGGGGGGS 165

RESULT 39
US-09-490-070A-178
Sequence 178, Application US/09490070A
Patent No. 6696248
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Ilag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon
APPLICANT: Plueckthun, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: Colin G. Sandercock, Esq. c/o Heller Ehrman
White & McAuliffe
STREET: 1666 K Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,070A
FILING DATE: 24-Jan-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Colin G. Sandercock, Esq.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 37629-0005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 912-2000
TELEFAX: (202) 912-2020
INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-09-490-070A-178

Query Match 94.6%; Score 106; DB 4; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGSGGGGGGGGGGGGS 20
|||||
Db 147 GGGSGGGGGGGGGGGGS 165

RESULT 40
US-09-490-153-178
Sequence 178, Application US/09490153
Patent No. 6706484
GENERAL INFORMATION:

APPLICANT: Knappik, Achim
Pack, Peter
Ilag, Vic
Ge, Liming
Moroney, Simon
Plusckthun, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,153
FILING DATE: 24-Jan-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 178:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-09-490-153-178

Query Match 94.6%; Score 106; DB 4; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GGS GGGGGGGGGGGGS 20
Db 147 GGS GGGGGGGGGGGGS 165

Search completed: August 19, 2005, 16:37:00
Job time : 43 secs